PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C07H 21/04, C12N 15/00 — A1

(11) International Publication Number: WO 87/ 02984

(43) International Publication Date: 21 May 1987 (21.05.87)

(21) International Application Number: PCT/US86/02440

(22) International Filing Date: 12 November 1986 (12.11.86)

(31) Priority Application Numbers: 797,421; 920,146

(32) Priority Dates: 12 November 1985 (12.11.85); 16 October 1986 (16.10.86)

(33) Priority Country: US

(71) Applicant: AMERICAN BIOGENETICS CORPORATION [US/US]; 19732 MacArthur Boulevard, Irvine, CA 92715 (US).

(72) Inventors: HATFIELD, G., Wesley ; 718 Goldenrod, Corona del Mar, CA 92625 (US). SHARP, Janice, A. ; 402 Fallingstar, Irvine, CA 92714 (US). MOYED, Harris, S. ; #14 Curie Court, Irvine, CA 92715 (US). LITTLE, Robert ; 6 Sage, Irvine, CA 92714 (US).

(74) Agents: SIMPSON, Andrew, H. et al.; Knobbe, Martens, Olson & Bear, 610 Newport Center Drive, Suite 1600, Newport Beach, CA 92660 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
With amended claims.

(54) Title: BIOGENETIC CASSETTE

(57) Abstract

A method for genetically engineering microbial strains which produce amino acids and other specialty chemicals comprising the molecular cloning of DNA chromosome fragments of the total genes involved in the biosynthesis of said chemicals and by genetic engineering, and/or mutation, removal of all biochemical and genetic regulation controlling the synthesis of such chemicals. Use may be made of a vector DNA molecule capable of providing amplification of the hybrid DNA construct containing the genes required for the biosynthesis of said chemicals. The resulting hybrid DNA molecule or biogenetic cassette is useful for transforming cells of an appropriate recipient strain capable of increased productivity of said chemicals. Also disclosed is a method of making a biogenetic cassette for catabolic pathways, including methanol utilization. Combination biogenetic cassettes, permitting, for example, the biosynthesis of a particular amino acid using methanol as a feedstock are also disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

BIOGENETIC CASSETTE

Background of the Invention

The present invention relates to the biotechnology industry and, more specifically, it relates to a method for preparing strains which produce excess quantities of amino acids or other chemicals, and for preparing strains which can metabolize such unconventional feedstocks as methanol.

Amino acids produced by microorganisms find extensive use as feedstuff and food additives in the agriculture and food industry, as components of various nutrient mixtures for medical purposes and as reagents in the chemical and pharmaceutical industries.

Methods are known for preparing strains which produce amino acids such as L-lysine, L-threonine, L-isoleucine and the like by using various mutagens, such as ultraviolet light, ionizing radiation, and chemical mutagens. The resulting mutant strains of microorganisms have specific genetically preconditioned defects in regulating metabolism and, due to such defects, they secrete specific amino acids into the nutrient medium. These mutant strains of microorganisms are identified and isolated by conventional methods based on the particular nutritive demand of a mutant (auxotrophy) or on resistance of a mutant to a structural analog of an amino acid that inhibits the growth of the parental strain (cf. British Pat. Nos. 1,258,380; 1,186,988; 1,316,888; Japanese Pat. No. 51-6237).

A method is also known for preparing mutants which produce amino acids on the basis of simultaneous resistance against an antimetabolite (such as an amino acid analog) and a co-inhibitor (a particular amino acid); (cf. U.S. Patent No. 3,756,916).

In all the above-mentioned cases, mutant strains capable of producing amino acids are prepared by single or step-by-step induction of mutations in the genetic

structure (genome) of a parental strain evolving the amino acid.

Also known are methods for constructing strains which produce amino acids comprising combining: (a) DNA chromosome fragments containing gene(s) controlling the synthesis of a selected amino acid and having a mutation conferring resistance to an amino acid analog destroying the negative regulation of the synthesis of this amino acid, with (b) a vector DNA molecule, to form a hybrid DNA molecule, capable of providing amplification of the presumed gene and therefore increased resistance to the amino acid analog resulting in increased secretion of the amino acid into the nutrient medium. These methods are based on random "shotgun cloning" of previously selected mutations, lacking defined insertion in an exact genomic configuration of genes involved in the biosynthesis of the amino acid, into a vector DNA molecule and amplification of this recombinant DNA molecule in contrast to directed methods for the purposeful cloning of the mutant gene.

Most genetically engineered products are manufactured using glucose as the primary carbon source. Citric acid, penicillin, and ascorbic acid are just a few examples of microbially produced products using glucose as a feedstock. Glucose, however, is a relatively expensive feedstock.

In many instances, a major operating cost of production is that of the carbon source required. For example, microbially produced biopolymers, used as thickeners in the food and cosmetic industries or to enhance oil recovery, are currently produced using glucose as the major carbon source. Microbially produced biopolymers may also replace certain synthetic plastics in the near future. The competitiveness of these biomaterials is, to a large extent, dependent on the cost of the carbon source utilized.

Methanol represents an alternative to glucose as a carbon source for microorganisms. Glucose sold in the United States is presently several times as expensive as methanol on a per pound basis. There is currently a billion gallon per year surplus capacity for producing methanol in the world. This surplus could grow to over two billion gallons per year in the near future. Although methanol can be made from crude oil, it is most easily made from natural gas and coal. As a feedstock, methanol is readily available, inexpensive, and likely to have greater long-term price stability than feedstocks based on crude oil. Many chemical manufacturers, in recognition of the competitive advantages of methanol as a feedstock, have developed new production facilities using it as the major source of carbon in many chemical processes.

Methylotrophic bacteria and yeasts, microorganisms which can obtain energy and carbon for growth from methanol, exist in nature and have been commercially exploited. For example, industrial-scale conversion of methanol to single-cell protein, to be used as a replacement for fishmeal and soybean meal in poultry feedstuffs, has already begun. This process uses a hexulose pathway organism, *Methylophilus methylotrophus*. Unfortunately, expression of hexulose pathway genes in *E. coli* has, to date, proved impossible. Nevertheless, the potential importance of methanol in commercial production is readily apparent.

Methanol has a host of properties that make it an ideal candidate as a feedstock for biotechnology fermentation processes. Methanol is extremely soluble in water. This eliminates the phase transfer problem associated with hydrocarbons. Unlike methane, methanol is not associated with a high risk of explosion. Methanol is readily purified. The metabolism of methanol by microorganisms requires less oxygen and generates less heat than does the metabolism of methane.

Accordingly, it is an object of the present invention to make possible the creation of organisms by genetic engineering techniques that have the ability to utilize methanol as a carbon source. Another object of the present invention is to provide a means whereby the methanol utilization ability of an engineered organism can readily be transferred to another organism.

Another object of the present invention is the creation of a biogenetic cassette for an unregulated biosynthetic pathway, such as amino acid biosynthesis.

Yet another object of the present invention is the creation of a combined biogenetic cassette for methanol utilization and amino acid biosynthesis.

In the context of the description of the present invention, the following expressions and terms are used:

The term "DNA" means deoxyribonucleic acid.

"Plasmids" are genetic elements reproducing in bacterial or yeasts cells irrespective of chromosomal DNA.

"Replication" means reproduction of genetic material.

"Transformation" and "transfection" mean transfer of genetic material to a host cell by means of isolated DNA.

The term "vector molecules" denotes DNA molecules of plasmids and phages ensuring transfer of a foreign genetic material to a host cell and its amplification.

The term "recombinant (hybrid) molecules" of DNA denotes a DNA molecule produced *in vitro* as a result of ligation, of different DNA molecules one of which is a vector molecule.

The term "clone" means genetically uniform progeny of one cell.

The term "cloning of genes" (molecular cloning) denotes preparation of clones of bacteria or yeasts containing said genes on a hybrid plasmid.

"Amplification" means increasing the number of genes in a cell.

"Operon" means a jointly controlled DNA unit containing a group of genes.

The term "catabolism" means the process of converting feedstuff compounds to intermediary metabolites.

The term "repression" means switching off transcription of genes or operons resulting in the inhibition of synthesis of gene products.

The term "repressor" denotes a regulating protein which stops the function of genes in combination with repression co-factors (usually final products of the biosynthetic or catabolic pathway or their derivatives.)

The term "biogenetic cassette" means a plasmid, bacteriophage, virus, or other form of self-propagating DNA constructed to contain all the genes, in the absence of negative regulatory features, genetic and metabolic, required for a complete biochemical pathway(s) for the catabolism of a specific feedstock or for the production of a specific amino acid or other chemical, or both, contained on a single hybrid DNA vector molecule. Where more than one enzyme (isozyme) can perform a given reaction in the pathway, the biogenetic cassette is constructed to contain a gene for only one of those enzymes, wherein the presence of that single gene and the absence of the other genes of similar function increase the production of the desired amino acid or chemical.

Brief Description of the Invention

The present invention relates to the production of microbial strains characterized by increased production of amino acids, or other chemicals. These microorganisms contain multiple copies of the genes necessary for biosynthesis of the desired end product, or utilization of the desired feedstocks, or both, but lack all or substantially all of the negative metabolic and genetic regulatory mechanisms limiting its synthesis.

The present invention also relates to the production of microbial strains having the ability to utilize methanol as a carbon source. These microorganisms may or may not contain multiple copies of genes necessary for utilization of methanol as a carbon source.

In accordance with one aspect of the present invention, a microorganism is prepared which produces an amino acid such as L-isoleucine. Fragments of *E. coli* chromosomal DNA containing all of the genes required for L-isoleucine biosynthesis are reconstructed in a configuration identical to their natural genomic configuration into a vector molecule of DNA, such as plasmid pBR322, to form a hybrid plasmid consisting of approximately 15 kilobases (kb) of DNA. The genes in this plasmid and their gene products are devoid of all negative metabolic and genetic regulation limiting L-isoleucine production.

Where more than one enzyme (isozyme) can perform a given reaction in the pathway, the biogenetic cassette may advantageously be constructed to contain a gene for only one of these isozymes. The presence of the single gene and the absence of the other isozymes in some instances may increase the production of the desired amino acid or chemical either by the nature of its metabolic or genetic regulation or by the nature of its substrate specificity. In the case of L-isoleucine and L-valine there exist three genes (*ilv*BN, *ilv*GM and *ilv*HI) for the acetohydroxy acid synthetase isozymes (AHASI, AHASII, AHASIII). A preferred biogenetic cassette for the production of L-isoleucine contains only the *ilv*GM genes for the production of AHASII (which according to its substrate specificity favors the synthesis of L-isoleucine) whereas the *ilv*BN and *ilv*HI genes code for the AHAS isozymes that favor the synthesis of L-valine.

Thus, in accordance with one aspect of the present invention, there is provided a recombinant DNA segment

containing genes coding for at least one enzyme for each step in the biosynthetic or metabolic pathway for a desired biological product, wherein the pathway has at least two steps, and preferably at least three steps, and wherein *in vivo* transcription of at least one of the genes is ordinarily subject to regulation and wherein at least one of the genes codes for an enzyme that is ordinarily subject to feedback inhibition, wherein all transcriptional regulatory sequences have been removed from the segment, and wherein each gene in the segment that would otherwise code for a feedback-inhibited enzyme has been modified so that feedback inhibition has been removed, so that upon transcription of the segment and translation of the resulting RNA, the biosynthetic or metabolic pathway is unregulated. Where the biosynthetic pathway is a pathway for an amino acid, preferred products include L-isoleucine, L-valine, L-leucine, L-tryptophan, L-histidine, L-threonine, and L-phenylalanine.

Where appropriate in light of the genetic regulatory mechanism, it is preferred that the genes in the segment are in natural chromosomal order.

One particularly preferred biogenetic cassette for L-isoleucine biosynthesis includes $ilvpe^{d}a^{d}G^{+}MEDA^{r}$ and *ilv*YC, wherein p indicates a promoter sequence, e indicates a leader sequence, a indicates an attenuator sequence, $^{+}$ indicates activity, ^{d} indicates deletion, and ^{r} indicates resistance to feedback inhibition.

In accordance with another preferred embodiment of the present invention, the DNA segment or biogenetic cassette prepared in accordance with this invention is contained in a recombinant DNA transfer vector. Preferred vectors include phages, other viral vectors, and plasmids. The plasmid is preferably a multicopy plasmid.

In accordance with yet another preferred embodiment of the present invention, the biogenetic cassette includes genes for both a biosynthetic pathway and a catabolic

pathway. One preferred catabolic pathway for methanol utilization requires a methanol oxidase encoding gene and a dihydroxyacetone synthase encoding gene.

In this preferred embodiment, transcriptional regulation may be removed by providing a promoter sequence for the methanol oxidase and dihydroxyacetone synthase encoding genes upstream from the methanol oxidase and the dihydroxyacetone synthase encoding genes, that is not subject to negative transcriptional regulation and is different from the wild-type promoter for the methanol oxidase encoding gene. The biogenetic cassette also preferably includes ribosome binding sites and a transcription terminator for the methanol oxidase and dihydroxyacetone synthase encoding genes.

Still another preferred embodiment of the present invention is a combined biosynthetic and catabolic biogenetic cassette for lactose utilization, which includes a *lac* operon in which the wild-type promoter sequence has been replaced with a different promoter sequence. The *lac* repressor gene, *laci*, has preferably been removed from the *lac* operon by site-specific deletion.

Another embodiment of the present invention is a biogenetic cassette consisting of a recombinant DNA segment comprising a methanol oxidase encoding gene and a dihydroxyacetone synthase encoding gene, both of which genes have been obtained from a methylotropic organism, and a promoter sequence directing the transcription of the genes, wherein the promoter sequence is different from the promoter sequence for the genes in the organism from which those genes came.

The present invention also includes a method for assembling an unregulated biogenetic cassette containing DNA coding for a biosynthetic or metabolic pathway having at least two, and preferably three steps, that is ordinarily regulated by both transcriptional regulation

and feedback inhibition, and wherein the cassette includes at least one gene for each step of the pathway. This method comprises the steps of obtaining a modified gene (such as a mutated or synthetic gene) for each enzyme in the pathway that is ordinarily subject to feedback inhibition, wherein the modification in each modified gene destroys the feedback inhibition, removing the DNA sequence or sequences responsible for negative transcriptional regulation of genes for the pathway to obtain a deattenuated or derepressed operon, assembling all of the genes coding for at least one of the enzymes for each step in the pathway, including the modified genes, into a single DNA segment, and cloning the resulting DNA segment. This method may also include the steps of removing a promoter sequence and inserting a different promoter sequence into the segment and inserting the segment into a plasmid and amplifying the segment in the plasmid in a microorganism. In a preferred embodiment, the product of the pathway is an amino acid, such as L-isoleucine, L-valine, L-leucine, L-tryptophan, L-histidine, L-threonine, or L-phenylalanine. Other amino acids may also be selected.

One particularly preferred method for making a biogenetic cassette for the L-isoleucine biosynthetic pathway includes the steps of removing the attenuator sequence from an *ilv*GMEDA operon to form a deattenuated operon, replacing the wild-type *ilv*A with L-isoleucine-resistant *ilv*A^r, and assembling together in a single polymeric DNA molecule the deattenuated operon containing the *ilv*A^r and *ilv*YC to form a biogenetic cassette coding for the biosynthesis of L-isoleucine.

The methods of the present invention optionally include the step of incorporating into a biosynthetic biogenetic cassette genes coding for a catabolic pathway, such as a methanol oxidase encoding gene and a dihydroxyacetone synthase encoding gene, or a *lac*

operon. Of course, the methods also may include obtaining expression of the biogenetic cassettes in a microorganism.

Further, the present invention comprehends a method for creating a biogenetic cassette for methanol utilization, comprising the steps of obtaining methanol oxidase and dihydroxyacetone synthase encoding genes from a methylotropic organism, removing the wild-type promoter sequence from the genes, creating a DNA segment containing the genes and a different promoter for the genes, and inserting the segment into a recombinant DNA transfer vector, such as a virus or plasmid.

Brief Description of the Drawings

Figure 1 is a diagram of the biosynthetic pathway for synthesizing L-isoleucine.

Figures 2A and 2B diagram a cloning strategy for assembling the *ilv* genes into a deregulated biogenetic cassette for the synthesis of L-isoleucine. DNA sequences of genes are represented by bold lines above the gene letter, circles represent DNA vector sequences. Abbreviations for restriction endonuclease sites are: H, HindIII; S, SmaI; P, PvuII; K, KpnI; Sa, SalI; X, XhoI; B, BglII.

Figure 3 diagrams an alternative cloning strategy and steps for forming a deregulated biogenetic cassette for the synthesis of L-isoleucine. DNA sequences of genes are represented by bold lines above the gene letter, circles represent vector sequences. Abbreviations are: H, HindIII; Sa, SalI; Su, Sau3IIA; T, TaqI; M, MnlI; Sp, SphI; Bg, BglII; Hi, HincII; X, XhoI; B, BamHI; R, EcoRI.

Figures 4A and 4B illustrate the cloning strategy and steps for forming a biogenetic cassette for methanol utilization. DNA sequences of genes are represented by bold lines, circles represent DNA vector sequences. Abbreviations for restriction endonuclease sites are: B, BamHI; H, HindIII; P, PvuII; A, AluI; Hi, HinfI; Sa,

Sau3A; R, EcoRII; E, EcoRI; T, TaqI; Sp, SphI; Bs, BstEII; Ba, BalI; Bg; BglII; N, NdeI; Sm, SmaI; X, XbaI.

Detailed Description of the Preferred Embodiments

The present invention employs methods of *in vitro* preparation of hybrid DNA molecules capable of replication and amplification, followed by introduction of these molecules into a recipient strain, by means of transformation or transfection. The vector molecules used are plasmid DNA or DNA of temperate bacteriophages, viruses, or other self-propagating DNA. Many of these methods are described in detail in the following publications:

1. Cohen S.N., Chang A.C.Y., Boyer H.W. and Helling R.B. (1973) *Proc. Nat. Acad. Sci. USA*, 70, 3240.

2. Green P.J., Betlach M.D., Boyer H.W., and Goodman H.N. (1974) *Methods in Molecular Biology*, 7, 87.

3. Tanaka T., and Weisblum B. (1975) *J. Bacteriol*, **121**, 354.

4. Clarke L., and Carbon J. (1975) *Proc. Nat. Acad. Sci. USA*, 72, 4361.

5. Bolivar F., Rodrigues R.L., Green P.J., Betlach M.C., Heyneker H.L., and Boyer H.W. (1977) *Gene*, 2, 95.

6. Kozlov J.I., Kalinina N.A., Gening L.V., Rebentish B.A., Strongin A.J., Bogush V.G., and Debabov V.G. (1977) *Molec. Gen. Genetics*, **150**, 211.

The biosynthesis of branched chain amino acids isoleucine and valine in all organisms is accomplished by a complex biochemical and genetic system and in *E. coli* comprises four separate operons and eleven genes (Fig. 1). The regulation of the synthesis of these amino acids is affected by several processes, including attenuation (*ilv*GMEDA and *ilv*BN), repression (*ilv*HI, *ilv*Y), induction (*ilv*C), catabolite repression (*ilv*BN), end product inhibition (threonine deaminase by L-isoleucine, acetohydroxy synthase I and III by L-valine, isopropylmalate synthetase by L-leucine), and end product

activation (threonine deaminase by L-valine). For a recent review of amino acid synthesis and regulation, see Herrmann, K. and Summerville, R.L., eds., Amino Acid, Biosynthesis and Genetic Regulation (1983).

In one biosynthetic pathway for the production of L-isoleucine, the starting material is L-threonine. This pathway is shown in Figure 1. Threonine deaminase (coded for by the *ilv*A gene) removes the amino group and a molecule of H_2O to form alpha-ketobutyrate. Next, acetohydroxy acid synthetase II (*ilv*GM), condenses that compound with pyruvate to form alpha-aceto-beta-hydroxybutyrate. In the next step, acetohydroxy acid isomeroreductase (*ilv*C) converts this product to alpha,beta-dihydroxy-alpha-methylvalerate, which in turn is dehydrated by dihydroxy acid dehydratase (*ilv*D) to make alpha-keto-beta-methylvalerate. This product is converted to L-isoleucine by an amino acid aminotransferase coded for by *ilv*E, all as shown in Figure 1.

As mentioned above, transcription of the *ilv*GMEDA operon is controlled by an attenuator sequence. Acetohydroxy acid synthetase I and III are subject to negative feedback regulation by L-valine. Wild-type *E. coli* K12 is acetohydroxy acid synthetase II negative (*ilv*G$^-$). Threonine deaminase is similarly subject to feedback inhibition by L-isoleucine. Transcription of *ilv*C is induced by its substrates acetohydroxybutyrate and acetolactate, by complexing with a positive activator (the product of the *ilv*Y gene).

Similarly, in other biosynthetic and catabolic pathways, repression, attenuation, positive activation and feedback inhibition of enzymes and other genetic and metabolic controls play an important role. For example, one catabolic operon, the *lac* operon of *E. coli* K12 responsible for the dissimilation of lactose, is genetically regulated by both repression and positive activation. In the absence of lactose in the nutrient

medium, the *lac* repressor (the product of the *lac*I gene) binds to the operator site in the promoter-operator region and prevents the transcription of the three genes of the *lac* operon by RNA polymerase. In the presence of lactose in the nutrient medium, a form of the lactose molecule complexes with the *lac* repressor in such a way as to prevent it from binding to the *lac* promoter-operator region, thus allowing transcription of the *lac* operon by RNA polymerase.

The *lac* operon is also regulated by catabolite repression, as are many other catabolic operons such as the arabinose operon, the ribose operon, etc. This regulation results from the fact that the transcription of these operons by RNA polymerase is activated by the binding of cyclic AMP receptor protein:cyclic AMP complex to the promoter region of these operons. In the presence of glucose in the nutrient medium, intracellular cyclic AMP levels are low, no CRP:cyclic AMP complex forms, and expression of these operons remains low (catabolite repression). In the absence of glucose and in the presence of other carbon sources (e.g., lactose, arabinose, ribose, etc.) cyclic AMP levels are high, CRP:cyclic AMP complexes form and the transcription of these operons is activated.

Another example of a biosynthetic operon is the *trp* operon of *E. coli* K12. This operon is responsible for the biosynthesis of L-tryptophan and is genetically regulated by both repression (like the *lac* operon) and attenuation as well as by metabolic end product inhibition (like the *ilv* biosynthetic pathway).

Briefly stated, attenuation is the regulation of transcription termination at a site (the attenuator) preceding the structural gene(s) of an operon. Translation of the leader RNA preceding the attenuator determines whether a termination structure (G + C stem-loop followed by a string of U residues) is formed in the

RNA transcript to signal transcription termination or is pre-empted to allow read-through transcription into the structural genes. Under conditions where the intracellular concentration of the appropriate amino acid(s) or other end product(s) is sufficient, translation of the leader RNA favors the formation of the termination structure and RNA polymerase terminates transcription at the attenuator. Alternatively, if the amino acid(s) or other product(s) become limiting, the ribosomes stall in the leader region, the termination signal is pre-empted, and RNA polymerase continues transcription into the structural genes.

Further examples of gene regulatory mechanisms are numerous and include such examples as regulation by anti-sense RNA (the *ompF* gene of *E. coli* K12), antitermination of transcription (N gene product regulation of late gene expression in bacteriophage lambda) and regulatory protein binding to mRNA to control translation (the ribosome operons of *E. coli* K12) and, in eukaryotes, specific DNA sequences (enhancers) are often required for efficient gene expression.

In all of the above examples these gene regulatory mechanisms are specified by unique and discrete DNA sequences. In the present invention we precisely delete, mutate, or alter by recombinant DNA methods these specific DNA sequences to remove negative regulation.

In assembling the biogenetic cassette of the present invention, all of the genes necessary for the biosynthesis of the desired end product or metabolism of a desired substrate are assembled together, in chromosomal order where appropriate (as discussd in Section D below). The metabolic and negative transcriptional regulatory mechanisms present in wild-type organisms have been deleted, modified, mutated, or otherwise defeated in the biogenetic cassette. In the L-isoleucine cassette, an active *ilvG*$^+$ replaces the wild-type inactive *ilvG*$^-$ in *E.*

coli K12; the wild-type *ilv*A has been replaced by a feedback-resistant *ilv*A^r; the inducer *ilv*Y is included with *ilv*C, and the *ilv*GMEDA attenuator is deleted, thereby producing a biogenetic cassette for the unregulated synthesis of L-isoleucine.

The specific steps necessary in assembling the biogenetic cassette will now be discussed in detail.

A. Removal of Allosteric Regulation

Genes for feedback-resistant enzymes may be obtained from microorganisms that have undergone either spontaneous or induced mutagenesis. Mutations are induced, for example, through the use of ionizing radiation, such as X-rays or ultraviolet radiation; chemical mutagens, such as nitrous acid, hydroxyl amine, ethyl methane sulfonate, and N-methyl-N'-nitro-N-nitrosoguanidine, and the like.

The mutation in the *ilv*A gene which confers resistance to feedback inhibition by L-isoleucine on threonine deaminase is a spontaneous mutation originally identified by the exhibition of a higher than normal intracellular concentration of threonine deaminase. Further characterization of this mutant strain demonstrates that the mutation is located in the *ilv*A gene and that the *ilv*A gene product of this strain, threonine deaminase, is highly resistant to inhibition by L-isoleucine. (Calhoun, D.H., Kuskee, J.S. and Hatfield, G.W. (1974) *J. Biol. Chem.* 250, 127-131.

B. Removal of Attenuator Sequence and Other Transcriptional Regulatory Site

The attenuator sequence may be removed by two possible methods. In the first, short segments of the promoter region are isolated as specific restriction endonuclease fragments and religated in such a manner as to reconstitute a fragment retaining the functionality of transcription initiation but for which the attenuator has

been removed. Alternatively, the attenuator sequence may be removed by site-directed mutagenesis generally as described by R.B. Wallace, *et al.*, (1980) *Science* 209, 1396-1400. In this method, a DNA fragment containing the attenuator region is cloned into the replicative form of the single stranded bacteriophage M13. A synthetic oligonucleotide, with sequences complementary to approximately 7 bases flanking the attenuator site to be deleted, is hybridized to the single stranded phage DNA and is used as a primer to direct the synthesis of DNA on this single-stranded circular DNA template. The synthetic oligonucleotide is thereby incorporated into the newly synthesized circular DNA molecule and the intervening attenuator sequences are eliminated in the newly synthesized strand of the heteroduplex DNA molecule. Transformation into *E. coli* followed by DNA replication resolves this heteroduplex into mutant (attenuator deleted) and parental DNA sequences. The synthetic oligonucleotide used to synthesize the mutation may also be used in either a positive or negative hybridization assay to screen for clones containing the attenuator sequence deletion.

C. Assembly of Operons and Genes in Chromosomal Order

The techniques used in assembling the genes into the biogenetic cassette are classical genetic engineering techniques, utilizing appropriate restriction endonucleases and ligases and cloning vectors. The strategy for assemblying the genes in a specified order into the cloning vector are unique for each set of genes to be cloned depending upon the position of unique restriction sites in the genes and in the cloning vector, for example, two strategies for assembling the genes of the L-isoleucine biogenetic cassette into the cloning vector pBR322 are depicted in Figures 2A and 2B and Figure 3 and described in detail in section E.

In the case of the L-isoleucine biogenetic cassette it is necessary to reconstruct exactly the natural chromosomal order of the *ilv*GMEDA and *ilv*YC operons. This is due to the fact that the expression of the distal gene of the *ilv*GMEDA operon, *ilv*A, may be regulated by opposing transcription from the autoregulated *ilv*Y gene. This regulation would be necessary to coordinate the flow of carbon from alpha-keto-butyrate and pyruvate through the L-isoleucine-L-valine biosynthetic pathway.

D. Introduction of Promoter Sequences

Different organisms, of course, recognize different promoter sequences. In accordance with the present invention, when the desired deregulated biogenetic cassette has been assembled, means are provided which facilitate removing the *E. coli* promoter sequences and introducing various new promoter sequences appropriate for different host bacteria or yeasts onto the biogenetic cassette into a unique site in the vector and for introducing the biogenetic cassette into a variety of cloning vectors. For example, in the case of the L-isoleucine cassette, the new promoter may be introduced via a BamHI site, derived during construction of pABC115 from the polylinker region of the vector pUC19, which is located where the attenuator deletion has occurred. This BamHI site in the plasmid is useful for inserting any of a variety of appropriate host bacteria or yeast promoter sequences, such as, for example, those set forth in Table 1 and many others too numerous to list.

TABLE 1

Host Bacterium or Yeast	Promoter	Reference
Corynebacterium	Corynebacteriophage	Tweten, R.K and

diptheriae	Beta toxin gene	Collier, R.J. *J. Bacteriol.* (1983) **156**, 680-685
	Corynebacteriophage Beta constitutive toxin gene (eg.tox-201)	Welkos, S.L. and Holmes, R.K. *J. Virol.*(1981) **37**, 946-954
	Corynebacteriophage Gamma c toxin gene	Michel, J.L.,*et al.* *J. Virol* (1982) **42**, 510-518
Bacillus subtilis	Sporulation gene spo 11A	Ydukin, M.D., *et al.* *J. Gen. Microbiol*, (1985) **131**, 959-962
	dnaE gene	Wang, L.F., *et al.* *J. Biol. Chem.* (1985) **260**, 3368-3372
	Aspartokinase 11 gene	Bondaryk, R.P. and Paulus, H. *J. Biol. Chem.* (1985) **260**, 585-591
	Subtilisin E gene	Wang, S.L., *et al.* *Proc. Natl. Acad. Sci. USA* (1984) **81**, 1184-1188
	rrnB operon	Steward, G.C. and Bott, K.F. *Nucl.*

		Acids Res. (1983) **11**, 6289-6300
	Alpha amylase genes (amyR2, amyE)	Yamazaki, H., *et al., J. Bacteriol* (1983) **156**, 327-337
	Glutamine amido-transferase gene	Makaroff, C.A. *et al., J. Biol. Chem.* **258** 10586-10593
Bacillus liceniformus	ermD gene	Gryczan,T., *et al. Mol. Gen. Genet.* (1984) **194**, 349-356
	sporulation gene	Ramakrishna, N., *et al. Nucl. Acids Res.* (1984) **12**, 1779-1790
	PenP gene	McLaughlin, J.R. *Nucl. Acids Res.* (1982) **10**, 3905-3919
Bacillus thuringeinsis	crystal protein gene	Wong, H.C., *et al. J. Biol. Chem* (1983) 258, 1960-1967
Pseudomonas sp.	Carboxypeptidase	Minton, N.P., *et*

		al.
	G2 gene	*Gene* (1984) 31, 31-38
Hansenula polymorpha	Methanol Oxidase gene	Ledeboer, A.M., *et al. Nucl. Acids Res.* (1985) **13**, 3063-3082
	Dihydroxyacetone Synthase gene	Janowicz, Z.A. *et al. Nucl. Acids Res.* (1985) **13**, 1043-3062
Saccharomyces cerevisiae	GAL1-GAL10 gene promoters	Johnston, M. and Davis, R.W. *Mol. and Cell Biol.* (1984) 4, 1440-1448
	HIS3 gene promoter	Struhl, K. *Proc. Natl. Acad. Sci. USA* (1982) 79, 7385-7389
	HIS4 gene promoter	Donahue, T.F., *et al., Cell* (1983) **32**, 89-98
	TRP5 gene promoter	Zalkin, H. and Yanofsky, C., *J. Biol. Chem.* (1982) **257**, 1491-1500
Saccharomyces	CYCc	Russel, P.R.

pombe	gene promoter	*Nature* (1983) **301**, 167-169
	ADH gene promoter	Russel, P.R. *Nature* (1983) **301**, 167-169

For example, when using the organism *Hansenula polymorpha*, the promoter sequence may advantageously be that of the methanol oxidase gene of this organism:

ATTCTATGAGGCCATCTCGACGGTGTTCCTTGAGTGCGTACTCCACTCTGTAGCGACT
GGACATCTCGAGACTGGGCTTGCTGTGCTGGATGCACCAATTAATTGTTGCCGCATGC
ATCCTTGCACCGCAAGTTTTTAAAACCCACTCGCTTTAGCCGTCGCGTAAAACTTGTG
AATCTGGCAACTCAGGGGGTTCTGCAGCCGCAACCGAACTTTTCGCTTCGAGGACGCA
GCTGGATGGTGTCATGTGAGGCTCTGTTTGCTGGCGTAGCCTACAACGTGACCTTGCC
TAACCGGACGGCGCTACCCACTGCTGTCTGTGCCTGCTACCAGAAAATCACCAGAGCA
GCAGAGGGCCGATGTGGCAACTGGTGGGGTGTCGGACAGGCTGTTTCTCCACAGTGCA
AATGCGGGTGAACCGGCCAGAAAGTAAATTCTTATGCTACCGTGCAGCGACTCCGACA
TCCCCAGTTTTTGCCCTACTTGATCACAGATGGGGTCAGCGCTGCCGCTAAGTGTACC
CAACCGTCCCCACACGGTCCATCTATAAATACTGCTGCCAGTGCACGGTGGTGACATC
AATCTAAA.

As another example, for *Eschericia coli* the *ilv*P_1P_2 promoter sequence of the *ilv*GMEDA operon may be advantageously used:

AAATTGAATTTTTTTCACTCACTATTTTATTTTTAAAAAACAACAATTTATATTGA
AATTATTAAACGCATCATAAAAATCGGCCAAAAAATATCTTGTACTATTTACAAA
ACCTATGGTAACTCTTTAGGCATTCCTTCGAACAAGATGCAAGAAAAGACAAA.

Likewise, any of the promoter sequences of the genes listed in Table 1 might be advantageously used for expressing the genes of the biogenetic cassette in host organisms for which they are particularly suited.

Other microorganisms that may be used as host

organisms for the present invention include bacteria such as *Clostridium sp.*, *Serratia sp.*, *Enterobacter sp.*, *Salmonella sp.*, *Klebsiella sp.*, *Rizhobium sp.*, *Rhodopseudomonas sp.* (and other photosynthetic bacterial species), *Xanthamonas sp.*, and the various methylotropic bacterial species, etc, and yeast such as *Candida sp.*, other *Saccharomyces sp.*, *Hansenula sp.*, *Mucor sp.* (and other filamentous fungi), etc.

Moreover, since sufficient information is now available for the introduction and expression of foreign genes into plant and animal cells and even into mature plants and animals, it is possible to use the technology of the present invention to insert biogenetic cassettes into these systems.

E. Selection of Cloning Vector

Different plasmids suitable as cloning vectors for construction of biogenetic cassettes have different host organism compatabilities due, primarily, to their ability to adapt to the DNA replication machinery of a given host organism. Consequently, in choosing a DNA vector for the construction and expression of the genes of a biogenetic cassette, in a given host organism, it is necessary to choose a vector capable of replication in that host organism. These DNA vectors also contain antibiotic resistance or other phenotypically selectable genes, compatible with growth and physiological properties of the host organism, in order to propogate, amplify and maintain the biogenetic cassette. A large number of suitable plasmids are known and are readily available. For any particular organism, a person of ordinary skill in the art will recognize what plasmids are the most suitable.

F. Other Considerations

In addition to efficient promotion of transcription and gene amplification for high level expression of a

biogenetic cassette the economical use of the cellular transcription apparatus is favored by transcription of only those sequences required for the proper efficient translation of a messenger RNA. That is, proper transcription termination, messenger RNA processing, and messenger RNA translation signals must be engineered into the biogenetic cassette and these signals must be compatible with the metabolic machinery of the host organism. Furthermore, it has been demonstrated that transcription termination is an important factor for maintaining high copy numbers of plasmids since transcription through the origin of replication antagonizes plasmid replication (Adams and Hatfield, (1984) *J. Biol. Chem*, **259**, 7399-7403). Examples of these signals and their proper utilization are described:

Transcription termination. In *E. coli* transcription termination signals are well defined. DNA sequences containing G + C rich inverted repeats followed by 6 or more Ts are known to be efficient transcription terminators. In examples which follow, the well characterized terminator of the *ilv*GMEDA attenuator will be used. Other well characterized, efficient transcription terminators could also be used. Transcription terminators in yeast are also well understood. In this case transcription termination is coupled to polyadenylation of the terminated mRNA transcript. An example of an efficient yeast transcription terminator is that of the *CYC*1 locus of *Saccharomyces cerevisea* (Zaret and Sherman, (1982) *Cell* 28, 563-573). In higher eukaryotic organisms the mechanism of transcription termination is less well understood; however, *in vitro* studies by Hatfield *et al.* ((1983) *Mol. Cell. Biol.*, 3, 1687-1693) have demonstrated that the bacteriophage lambda 4S terminator contains DNA sequences that are recognized by the eukaryotic RNA polymerase II, the polymerase that transcribes structural

genes.

Messenger RNA processing. Current evidence suggests that polyadenylation of eukaryotic messenger RNAs in eukaryotic organisms is important for message stability and for the transport of the mRNA from the nucleus to the cytoplasm. The signals for polyadenylation, AATAAA together with nearby uncharacterized flanking sequences, have been documented for many eukaryotic genes. These sequences are found in structural genes immediately downstream of the amino acid coding region.

Messenger RNA translation. In *E. coli* it has been documented by Shine, J. and Dalgarno, L. ((1975) *Nature* 254, 34-38) and many other investigators that a sequence complementary to the 3'-end of the 16S ribosomal RNA is located approximately 9-17 nucleotides prior to the translation initiation codon on the mRNA. In the examples that follow the well characterized translation initiation sequence for the leader polypeptide of the *ilv*GMEDA operon will be used. Any other well defined translation initiation sequence could also be used. Translation initiation is less well understood in eukaryotes; however, as a general rule, translation appears to initiate at the first initiation codon (Kozak, (1985) *Micro. Rev.*, 47, 1-45).

In addition to the large number of general and specific considerations set forth above, it must be recognized that the mechanisms of expression vary from organism to organism. Many of these mechanisms have been elucidated at this time. Many additional mechanisms are the subject of continuing research and will be understood more fully in the future. A person of ordinary skill in the art will understand, of course, what DNA sequences are needed to obtain expression of the biogenetic cassette of the present invention in any of the ever-expanding number of organisms in which expression of the biogenetic cassette may be desired.

By use of the foregoing strategy and the steps discussed therein, the present invention can be applied to a wide range of biosynthetic or metabolic and catabolic pathways for production and/or utilization of both natural and altered biological products. The applicability of these steps is not limited to any particular type of biosynthetic or catabolic pathway, since the regulatory mechanisms and the genetics involved relate to and are applicable to an extremely broad range of natural systems. Thus, the following examples are not limiting, but rather are representative of the diversity of the pathways for which a biogenetic cassette can be constructed.

The present invention will now be described in great detail in the context of several biosynthetic and catabolic pathways. As has been discussed, there are a number of regulatory mechanisms involved in the biosynthesis of amino acids. Perhaps the most complex pathway, from a regulatory standpoint, is that for L-isoleucine.

Example I: L-isoleucine Biogenetic cassette

We describe here the construction of a biogenetic cassette containing all of the genes required for the biosynthesis of L-isoleucine (*ilv*GMEDA, *ilv*YC) in their natural chromosomal orientation except having the *ilv*GMEDA attenuator removed by genetic engineering methods and having the wild type *ilv*A gene replaced by a gene whose product (threonine deaminase) is resistant to end product inhibition by L-isoleucine because of mutation. All of the genes required for L-isoleucine biosynthesis are obtained from the *E. coli* DNA chromosome by isolation of appropriate chromosome restriction endonuclease fragments in appropriate DNA vectors. The two alternative cloning strategies followed in this example are illustated in Figures 2A and 2B and Figure 3.

Construction of Clones for the Assembly of Biogenetic cassette

A 4.7 Kilobase (Kb) HindIII fragment containing *ilv*peaG$^-$ME' (p=promoter, e=leader, a=attenuator) is cloned into the HindIII site of plasmid pBR322 to yield plasmid pABC5. Plasmid pABC5 is described in the literature by Lawther, *et al.*, (1979) *Nucl. Acids Res.* 7,2289-2301, where it is given the designation pRL5. The preparation of this plasmid is described in detail in this article. A 4.8 Kb HindIII-BglII fragment containing *ilv*'EDAY' is used to replace the HindIII-BamHI fragment of pABC5 to yield pABC55 (*ilv*peaG-MEDAY'$^-$9.5 Kb). The active *ilv*G gene is obtained by intracellular genetic recombination by growing pABC5 for several generations in an L-valine resistant *E. coli* K12 strain to yield pABC1 (*ilv*peaG$^+$ME').

The plasmid pABC1 has been described by Lawther *et al.*, (1981) *Proc. Natl. Acad. Sci. USA* **78**, 922-925, where it is identified as pRL101. This article describes in detail the preparation of pRL101.

Plasmid pABC55 is maintained by the American Type Culture Collection (ATCC) in Rockville, Maryland as a Budapest Treaty deposit under accession number 67,228, and plasmid pABC1 is maintained by ATCC as a Budapest Treaty deposit under accession number 67,230.

The SmaI-HindIII fragment from pABC1 is inserted into the SmaI-HindIII sites of the replicative form of bacteriophage mp18 to yield the hybrid mABC102. The 32 bases of the altenuator are excised from mABC102 by the oligonucleotide site directed deletion method of R.B. Wallace *et al.*, ((1980) *Science* **209**, 1396-1400) to yield mABC102d.

In an alternative method for deleting the attenuator of the *ilv* promoter, a 384 bp SauIIIA-TaqI fragment is isolated from pABC55, filled in with the Klenow fragment of *E. coli* DNA polymerase I, and cloned into the SmaI site

of the vector pUC19 to form pJGLP$_1$P$_2$. Subsequently an 800 bp MnlI-SphI fragment extending from immediately downstream of the attenuator region into the proximal portion of *ilv*G is isolated from pABC1 and cloned into HincII-SphI digested pJGLP$_1$P$_2$ to form pP$_1$P$_2$*att*. This construction results in the incorporation of a BamHI site from the polylinker of pUC19 between the P$_1$P$_2$ promoters and the beginning of the *ilv*G gene for use in inserting alternative promoters, as described in Section D, supra.

An XhoI fragment (4.4 Kb fragment located at approximately 84 min. on the *E. coli* chromosome) isolated from an *E. coli* K12 strain (CU18) containing a spontaneous mutation in threonine deaminase resulting in resistance to feedback inhibition by L-isoleucine, (Calhoun, D.H., Kuska, J.S. and Hatfield, G.W. (1974) *J. Biol. Chem.* 250, 127-131) is inserted into the unique XhoI site of pIC19R (Marsh, J.L., *et al.* (1984) *Gene* 32, 481-485) in an orientation such that the threonine deaminase gene *ilv*A is transcribed from the *lac* promoter to yield pABC101 (*ilv*A^{r}YC). (CU18 is maintained by ATCC as a Budapest Treaty deposit under accession number 53,549.) This clone is selected and isolated by its ability to complement an *ilv*A$^-$ *E. coli* strain CU623, described by Smith *et al.* (1979) *Mol. Gen. Genet.* 169, 229-314.

Alternatively, any fragment from CU18 on which the feedback resistance mutation in *ilv*A is included could be ligated into any convenient vector and isolated by plasmid sizing and restriction enzyme analysis. For example, an alternative method is to clone a 3.9Kb SalI-XhoI fragment of CU18 into SalI-XhoI digested pIC19R to form pABC102.

Construction of L-isoleucine Biogenetic Cassette: Alternative 1:

The L-isoleucine producing cassette for *E. coli* is assembled according to following steps, as ilustrated in Figures 2A and 2b:

1. A SmaI-HindIII DNA fragment is isolated from pABC5 and ligated into the PvuII-HindIII sites of pBR322 to yield pABC103 which is screened and isolated by plasmid size and restriction analysis (note PvuII sites and SmaI are sacrificed).

2. The resultant single PvuII fragment of pABC103 is replaced with the PvuII fragment from mABC102d to yield pABC104. This clone is selected by its ability to complement an AHAS$^-$ *E. coli* strain CU888, described by Baez, *et al.*, (1979) *Molec. Gen. Genet.* **169**, 289-297.

3. The SalI fragment of pABC55 is inserted into the SalI site of pABC104 to yield pABC105. pABC105 is selected by its ability to compliment *E. coli ilv*D$^-$ or *ilv*E$^-$ strains CU624, described by Smith, *et al.*, (1976) *Mol. Gen. Genet.* **148** 111-124; or CU692, described by Smith, *et al.* (1979) *Mol. Gen. Genet.*, **169**, 299-314.

4. The XhoI fragment from pABC101 is ligated into the XhoI site of pABC105 to yield pABC106 (*ilv*pe^{d}a^{d}G$^+$MEDA^{r}YC).

5. The cassette is made "mobile" by placing a linker in the deleted attenuator region by partial digestion of the plasmid with EcoRII, followed by insertion of an 8 base BamHI linker in this site. A BamHI linker is also inserted at the XhoI site external to the operon by partial digestion. These are the only BamHI sites in pABC106. The upstream BamHI site is also used for insertion of novel or known promoter sequences.

Construction of L-isoleucine Biogenetic Cassette: Alternative 2:

In an alternative construction, the cassette may be assembled as illustrated in Figure 3, thus:

1. A 3.0 Kb SalI fragment is isolated from pABC55 and ligated into SalI digested pABC1 which is screened for insertion of the SalI fragment in the appropriate

orientation by restriction enzyme analysis to yield pABC111.

2. The 1.2 Kb EcoRI-SphI attenuator deleted promoter fragment from pP_1P_2att is cloned into EcoRI-SphI digested pABC111 to form pABC114, whose integrity is confirmed by analytical restriction enzyme analysis.

3. The 3.9 Kb SalI-XhoI fragment from pABC102 is cloned into partial SalI-digested pABC114 to yield pABC115 whose identity is confirmed by restriction enzyme analysis.

4. A BamHI linker is added to a unique XmaIII site located distal to *ilv*YC in order to "mobilize" the cassette. This provides a BamHI site in the former attenuator region and another at the operon's distal insert-vector junction.

The completed cassette can now be expressed in *E. coli* or other suitable organisms. Unlike mutant organisms which produce an excess of L-isoleucine in combination with L-valine, the selectivity of the (AHAS) isozyme at the branchpoints in the pathway used in the present biogenetic cassette greatly decrease the amount of L-valine produced, with respect to L-isoleucine, thereby facilitating the purification process and reducing costs. Any one of the genes of this pathway might be further mutated to enhance the selectivity of L-isoleucine substrates.

Example II: L-valine Biogenetic cassette

The biosynthetic or metabolic pathway of L-valine synthesis differs from that of L-isoleucine in that the first step in L-valine biosynthesis involves the condensation of two pyruvate molecules, whereas, the analagous step in L-isoleucine biosynthesis involves the condensation of one molecule of pyruvate with one molecule of alpha-ketobutyrate. This enzymatic step is catalyzed in *E. coli* by three isozymes, AHASI (*ilv*BN), AHASII

(*ilv*GM), AHASIII (*ilv*HI). Of these isozymes, AHASII favors the synthesis of L-isoleucine whereas AHASI favors the synthesis of L-valine. Thus, the L-valine biogenetic cassette contains the *ilv*BN genes instead of the *ilv*GM genes. To remove the activity of the *ilv*GM gene and to avoid transcriptional polarity, pABC106 or pABC115 is digested at the unique SphI site in *ilv*G and deleted with Bal31 digestion, and religated. The correctly deleted plasmid is identified directly by plasmid sizing and restriction enzyme analysis of recombinant clones to create pABC108.

The *ilv*BN gene is inserted into the operon distal BamHI site of pABC108 by the following method. The altenuator of *ilv*BN is deleted from pCH4 either by site-directed oligoneucleotide mutagenesis after subcloning *ilv*BN into bacteriophage mp18 or by appropriate restriction fragment reconstruction of the promoter region, as described for the L-isoleucine biogenetic cassette. Plasmid pCH4 is described by Wek, R., Hauser, C.A. and Hatfield, G.W., (1985) *Nucl. Acids Res.* **13**, 3995-4010, and is maintained as a Budapest Treaty Deposit by ATCC under accession number 67,229. Into this is cloned a 1713 bp NaeI-SacI fragment from the L-valine resistant *ilv*BN operon from *E. Coli* strain MF2324 (Sutton, *et al.* (1981) *J. Bact.* **148**, 998-1001).

A 2.4 KB RsaI (on which has been ligated a BamHI linker) - BclI (partial digestion) fragment containing *ilv*$a^{d}B^{r}M$ is then inserted into the distal BamHI site of pABC108 to form pABC109.

The pABC109 plasmid is selected by its ability to complement an AHAS^{-} *E. coli* strain, CU888 or by directly sizing and restriction analysis. Since the distal BamHI site is sacrificed in the above construction, the L-valine cassette is mobile due to the fact that all of the genes are bounded by two unique BamHI sites.

Example III: A Biogenetic Cassette for Methanol Utilization by *E. coli*

A pathway which includes dihydroxyacetone as a key intermediate for the assimilation of methanol into the cell constituents of *E. coli* is established by the introduction of a biogenetic cassette containing the methanol oxidase (MOX) and dihydroxyacetone synthase (DHAS) of methanol-utilizing yeast into this organism. The preferred methanol-utilizing yeast, *Hansenula polymorpha*, American Type Culture Collection Catalog No. 34438, produces a methanol oxidase enzyme (MOX) which converts methanol to formaldehyde. The formaldehyde, in turn, is used by the dihydroxyacetone synthase enzyme (DHAS) to convert xyulose-5-phosphate to dihydroxyacetone and glyceraldehyde-3-phosphate. These products, in turn, enter the glycolytic pathway and are used in cellular metabolism. The cloning strategy used in this example is graphically illustrated in Figures 4A and 4B.

Construction of Clones for Assembly of the MOX cassette

The MOX gene is isolated as a 2.9 Kb PvuII-XbaI fragment of genomic DNA, and is ligated into SmaI-XbaI digested pUC18 to yield pABC201 whose identity is confirmed by restriction enzyme analysis and DNA sequence analyses of the 5' portion of the gene. The 3' untranslated region of the gene is deleted by digesting pABC201 with HindIII, limited digestion with Bal31, adding HindIII linkers, and cloning a 575 bp BstEII-HindIII fraction back into the BstEII-HindIII 4.5 Kb fragment from pABC201. A number of clones are identified by restriction enzyme analysis, and the DNA sequence of the 3' portion of the MOX gene in these is obtained to determine the end point of the Bal31 deletion. One of these, pABC210, which terminates 6 bp downstream from the UAA termination codon, is used for subsequent manipulations.

To delete the 5' untranslated region, a 1994 bp BalI-HindIII fragment from pABC210 is ligated to SmaI-HindIII pUC18, linear ligation products of 4.5 Kb are isolated, and ligation is continued in the presence of a synthetic NdeI linker having the sequence CCATATGG. In this manner, the 5' four nucleotides of the MOX coding region, which are deleted by the BalI digestion and include the initator methionine codon AUG required for translation, are reconstituted by the ATGG sequence in the synthetic oligomer. One of several appropriate clones identified by restriction enzyme analysis, pABC211, is then partially digested with NdeI, filled in with the Klenow fragment of *E. coli* DNA polymerase I, and ligated with BamHI linkers. A 2.0 Kb BamHI-HindIII fragment is then isolated and ligated into BamHI/HindIII-digested pUC18 to give pABC212.

The DHAS gene is isolated as a 3.2 Kb BamHI-HindIII fragment of genomic DNA, is inserted into BamHI-HindIII digested pUC18, and is identified by restriction enzyme analysis and DNA sequencing of the 5' end of the gene. This clone is designated as pABC204. A 521 bp AluI fragment is isolated, redigested with BglII, and the 358 bp BglII-AluI fragment is inserted back into the 5.4Kb BglII-BamHI fragment of pABC204 in which the BamHI site has been filled in with the Klenow fragment of *E. coli* DNA polymerase I. In this manner a BamHI site is maintained at the 3' end of the DHAS gene since the AluI/BamHI (Klenow) junction provides the full GGATCC recognition sequence. Several appropriate clones are identified, one of which, pABC214, is used in further manipulations.

To delete the 5' untranslated region of the DHAS gene, pABC204 is digested with HindIII, limit digested with Bal31, ligated with HindIII linkers and digested with BstEII and HindIII. A 380 bp BstEII-HindIII fraction is isolated and ligated into the 4.7 Kb HindIII-BstEII fragment of pABC204 and a number of clones are identified

by restriction enzyme analysis. The Bal31 deletion in one of these clones, pABC215, is shown by DNA sequence analysis to terminate at position -1 where the A of the ATG methionine initiation codon is designated as +1. The 380 bp HindIII-BstEII fragment from this clone is ligated to the BstEII-HindIII digested pABC214 to give pABC216.

For expression of MOX and DHAS in *E. coli*, *E. coli* promoter, terminator and ribosome binding sites are required. These are derived from the attenuator region of *ilv*. The terminator of the fragment is obtained as a 146 bp SauIIIA-HinfI fragment, filled in with the Klenow fragment of *E. coli* DNA polymerase I, and cloned into the HincII site of pUC19 to form pABC207. This sequence is transferred as a BamHI-SphI fragment from sites in the polylinker into BamHI-SphI digested pBR322 to form TpBR. The promoter and ribosome binding sites are obtained from an EcoRII 415 bp fragment filled in with Klenow and ligated into the SmaI site of pUC19 to form pABC209. In order to eliminate the NH_4-terminal amino acid residues of the *ilv* leader peptide, pABC209 is digested with BamHI, treated with Bal31, ligated with BamHI linkers, and digested with EcoR1 and BamHI. A 420 bp fraction is isolated and ligated into pBR322. A number of clones are isolated by restriction enzyme analysis and DNA sequencing, one of which (pABC219) has a Bal31 deletion end point 4 b.p. upstream of the methionine ATG initator codon of the *ilv* leader peptide. The 440 bp EcoR1-BamHI fragment of pABC219 is cloned into EcoR1-BamHI digested TpBR to give pABC220 in which a unique BamHI site is present between P_1P_2 and the terminator sequence.

Construction of the MOX Biogenetic Cassette.

1) The translated portions of the MOX and DHAS genes are joined by ligating HindIII digested pABC212 and pABC216, digesting the ligation mixture with BamHI, and isolating the 4.2 Kb BamHI MOX-DHAS fusion fragment which

is subsequently cloned into the BamHI site of pBkHindIII^d (pBK322 in which the HindIII site has been deleted) to give pABC217.

2) The ribosome binding site contained on a 30 b.p. Klenow treated TaqI-BamHI fragment from pABC219 is cloned into HindIII-digested, Klenow filled pABC217 to give pABC221, in which the appropriate orientation of the ribosome binding site is confirmed by DNA sequence analysis.

3) The BamHI 4.3 Kb MOX-rbs-DHAS fragment from pABC221 is cloned into BamHI-digested pABC220 to form pABC230.

Thus, a methanol utilizing operon complete with a promoter, two ribosome binding sites, the two structural genes required for methanol utilization and a transcription terminator are contained in the plasmid pABC230, facilitating the expression of these genes in *E. coli*.

The promoter for the MOX gene is ordinarily positively regulated by methanol and negatively regulated by glucose. Because the naturally-occurring MOX promoter has been replaced with the *ilv*GMEDA promoter, all known negative transcriptional regulation has been removed from the resulting biogenetic cassette.

Example IV: Biogenetic Cassette for the S10 Operon of *E. coli* K12

Previous studies have shown that ribosomal protein L4 specifically inhibits the expression of its own operon, the 11-gene S10 operon of *E. coli* K12 (Lindahl, L., *et al.* (1983) *Cell*, **33**, 241-248). In the presence of excess L4, transcription of the S10 operon is terminated 140 bases past the transcription initiation site. This attenuation of S10 operon expression by L4 ensures a constant but limited synthesis of operon products. The S10 operon is also transcriptionally regulated by a growth rate control

mechanism exerted at the level of transcription initiation. Using the same techniques as set forth in Examples 1-3 a biogenetic cassette for the deregulated expression of the products of the S10 operon is constructed by producing a DNA fragment containing the 11 genes of the operon with the attenuator site deleted (as in the construction of the L-isoleucine biogenetic cassette, Example I) and with a new, strong promoter sequence replacing the S10 promoter (e.g. *trp*P, *tac*P, or *ilv*P_1P_2) and inserting this fragment into an amplifiable plasmid vector. Since no metabolic controls are known for this operon this construction would represent a biogenetic cassette containing all 11 genes of the S10 operon, on a single DNA fragment, inserted into a single plasmid, completely devoid of all known negative regulation. In order to maintain the integrity of the biogenetic cassette on the plasmid, a recombination-deficient (*E. coli*) host should be used.

Example V: Biogenetic Cassette for Lactose Utilization

Using the same techniques as set forth in Examples 1-3, a genetic cassette for the deregulated expression of the *lac* operon (Jacob F. and Monod J. (1961) *J. Mol. Biol.* 3, 318-365) of *E. coli* K12 is prepared as follows: The *lac* I gene, which encodes the *lac* repressor, is removed by site specific deletion. A DNA fragment containing the *lac*Z, *lac*Y and *lac*A genes of the *lac* operon with the natural *lac* operator-promoter region replaced with a strong constitutive promoter sequence is placed into an amplifiable high copy plasmid vector. This genetic cassette containing all the genes of the *lac* operon devoid of negative regulation for the efficient catabolism of lactose can also be prepared (as in Example 1) with a unique restriction site(s) at the promoter site for the insertion of host specific promoter and vector sequences so that this cassette can be functional in any host for

which specific promoter and vector sequences are now, or later become, available.

Example VI: Catabolic and Biosynthetic Cassette

Example 1 and 2 describe biogenetic cassettes for the biosynthesis of the natural products, L-isoleucine and L-valine, respectively. Example 3 and 5 describe biogenetic cassettes for the catabolism of natural products, methanol and lactose, respectively. A genetic cassette for the conversion of a specific natural product into another specific natural product can be constructed by combining a catabolic operon with a biosynthetic operon in the same plasmid vector. In the above cases it is possible, for example, to construct biogenetic cassettes for the production of L-isoleucine or L-valine from either methanol or lactose. This same strategy can be employed to produce any number of natural products from a wide variety of starting materials in appropriate host organisms by combining the genes of any given catabolic and biosynthetic pathways in the same biogenetic cassette.

Example VII: L-leucine Biogenetic Cassette

The biosynthesis of L-leucine is regulated at the level of transcription by attenuation of the *leu*ABCD operon and by end-product inhibition of the first enzyme of the L-leucine pathway (alpha-isopropylmalate synthase, *leu*A). No other repression or induction mechanism is utilized by the *leu* operon. Using the same techniques as set forth in Examples 1-3 a biogenetic cassette for the deregulated expression of the *leu* operon of *E. coli* K12 is prepared as follows: The *leu*ABCD attenuator is removed by procedures similar to those described in Example 1 using site specific deletion. A DNA fragment, from 2 min. on the *E. coli* chromosome, containing the *leu*ABCD genes of the *leu* operon, in which the natural *leu* promoter region has been replaced with a strong constitutive promoter

sequence, is placed into an amplifiable high copy plasmid vector. The gene coding for alpha-isopropylmalate synthase (*leu*A) is replaced with a gene coding for a feed-back resistant alpha-isopropylmalate synthase. This biogenetic cassette, containing all of the genes of the *leu* operon devoid of negative regulation for the efficient biosynthesis of L-leucine can also be prepared (as in Example 1) with a unique restriction site(s) at the promoter site for the insertion of host specific promoter and vector sequences so that this cassette can be functional in any host for which specific promoter and vector sequences are now, or become, available.

Example VIV: L-tryptophan Biogenetic Cassette

The biosynthesis of L-tryptophan is regulated at the level of transcription by attenuation of the *trp*EDCBA operon as well as by repression of the operon by the *trp* repressor and by end-product inhibition of the first enzyme of the L-tryptophan pathway, anthrinilate synthetase (*trp*E). Using the same techniques as set forth in Examples 1-3, a biogenetic cassette for the deregulated expression of the *trp* operon of *E. coli* K12 is prepared as follows: The *trp* attenuator is removed by procedures similar to those describes in Example 1 using site specific deletion. A DNA fragment containing the *trp*EDCBA genes of the *trp* operon in which the natural *trp* repressor-promoter region has been replaced with a strong constitutive promoter sequence is placed into an amplifiable high copy plasmid vector. The gene coding for anthrinilate synthetase (*trp*E) is replaced with a feed-back resistant anthrinilate synthetase *trp*E gene. This biogenetic cassette, containing all of the genes of the *trp* operon devoid of negative regulation for the efficient biosynthesis of L-tryptophan, can also be prepared (as in Example 1) with a unique restriction site(s) at the promoter site for the insertion of host specific promoter

and vector sequences so that this cassette can be functional in any host for which specific promoter and vector sequences are now, or become, available.

Example IX: L-histidine Biogenetic Cassette

The biosynthesis of L-histidine is regulated at the level of transcription by attenuation of the *his*GDCBHAFIE operon and by end-product inhibition of the first enzyme of the L-histidine pathway, (*his*G). Using the same techniques as set forth in Examples 1-3, a biogenetic cassette for the deregulated expression of the *his* operon of *E. coli* K12 is prepared as follows: The *his*GDCBHAFIE attenuator is removed by procedures similar to those described in Example 1 using site specific deletion. A DNA fragment, from 44 min. on the *E. coli* chromosome, containing the *his*GDCBHAFIE genes of the *his* operon with the natural *his* attenuator-promoter region replaced with a strong constitutive promoter sequence is placed into an amplifiable high copy plasmid vector. The wild-type *his*G gene is replaced with a *his*G gene coding for a feed-back resistant *his*G gene product. This biogenetic cassette, containing all of the genes of the *his* operon devoid of negative regulation for the efficient biosynthesis of L-histidine, can also be prepared (as in Example 1) with a unique restriction site(s) at the promoter site for the insertion of host specific promoter and vector sequences so that this cassette can be functional in any host for which specific promoter and vector sequences are now, or later become, available.

Example X: L-threonine Biogenetic Cassette

The biosynthesis of L-threonine is regulated at the level of transcription by attenuation of the *thr*ABC operon and by end-product inhibition of the first enzyme of the threonine pathway, homoserine dehydrogenase (*thr*A). Using the same techniques as set forth in Examples 1-3, a

biogenetic cassette for the deregulated expression of the *thr* operon of *E. coli* K12 is prepared as follows: The *thr*ABC attenuator is removed by procedures similar to those described in Example 1 using site specific deletion. A DNA fragment, from 0 min. on the *E. coli* chromosome, containing the *thr*ABC genes of the *thr* operon, with the natural *thr* attenuator-promoter region replaced with a strong constitutive promoter sequence, is placed into an amplifiable high copy plasmid vector. The gene coding for homoserine dehydrogenase (*thr*A) is replaced with a feed-back resistant homoserine dehydrogenase. This biogenetic cassette containing all of the genes of the *thr* operon devoid of negative regulation for the efficient biosynthesis of L-threonine can also be prepared (as in Example 1) with a unique restriction site(s) at the promoter site for the insertion of host specific promoter and vector sequences so that this cassette can be functional in any host for which specific promoter and vector sequences are now, or later become, available.

Example XI: L-phenylalanine Biogenetic Cassette

The biogenetic cassettes described in Examples 1-10 involve biosynthetic or catabolic pathways that can be inserted in host organisms with only minor perturbations of tributory metabolic pathways. Many other examples of biogenetic cassettes can be envisioned which for maximum utilization and efficiency would require further genetic engineering of the host organism. For example, the biosynthesis of L-phenylalanine is regulated at the level of transcription by attenuation of the *phe* operon and by end-product inhibition of the first enzyme of the L-phenylalanine pathway. Using the same techniques as set forth in Examples 1-3, a biogenetic cassette for the deregulated expression of the *phe* operon of *E. coli* K12 is prepared as follows: The *phe* attenuator is removed by procedures similar to those described in Example 1 using

site specific deletion. A DNA fragment, from 56 min. on the *E. coli* chromosome, containing the genes of the *phe* operon with the natural *phe* attenuator-promoter region replaced with a strong constitutive promoter sequence, is placed into an amplifiable high copy plasmid vector. The gene coding for the feed-back resistant first enzyme of the *phe* pathway is replaced with a gene encoding a feed-back resistant enzyme. This biogenetic cassette containing all of the genes of the *phe* operon devoid of negative regulation for the efficient biosynthesis of L-phenylalanine can also be prepared (as in Example 1) with a unique restriction site or sites in the promoter site for the insertion of host specific promoter and vector sequences so that this cassette can be functional in any host for which specific promoter and vector sequences are now, or become, available. However, since L-phenylalanine is but one member of the aromatic amino acid family (L-phenylalanine, L-tryptophan and L-tyrosine) and since all of these amino acids are synthesized from a common branch point intermediate (chorismic acid) which is synthesized by a pathway subject to regulation by all of the aromatic amino acids, deregulation of this common pathway would also be important for the efficient production of L-phenylalanine. For example it would be advantageous in this case to replace one or two of the three genes for the feed-back sensitive DHAP synthase isozymes in the chorismic acid pathway with genes for feed-back insensitive DHAP synthases.

Although the invention has been described with reference to specific examples, many other embodiments fall within the scope and spirit of the present invention. Accordingly, it is intended that the protection afforded by this patent be limited only by the scope of the following claims, and reasonable equivalents thereof.

WHAT IS CLAIMED IS:

1. A recombinant DNA segment containing genes coding for at least one enzyme for each step in the biosynthetic or metabolic pathway for a desired biological product, wherein said pathway has at least two steps, and wherein *in vivo* transcription of at least one of said genes is ordinarily subject to regulation and wherein at least one of said genes codes for an enzyme that is ordinarily subject to feedback inhibition, wherein all transcriptional regulatory sequences have been inactivated or removed from said segment, and wherein each gene in said segment that would otherwise code for a feedback-inhibited enzyme has been modified so that feedback inhibition has been removed, so that upon transcription of said segment and translation of the resulting RNA, said biosynthetic or metabolic pathway is unregulated, said segment having a first restriction site and a second restriction site with said genes located between said first and second restriction sites, wherein said first and second restriction sites do not occur elsewhere in said segment.

2. The segment of Claim 1, coding for the biosynthetic pathway for an amino acid.

3. The segment of Claim 2, wherein said amino acid is L-isoleucine.

4. The segment of Claim 2, wherein said amino acid is L-valine.

5. The segment of any of Claims 1-4, wherein said genes are in natural chromosomal order.

6. The segment of Claim 3, wherein said segment comprises *ilv*$p^{d}a^{d}G^{+}MEDA^{r}YC$, wherein p indicates a promoter sequence, e indicates a leader sequence, a indicates an attenuator sequence, $^{+}$ indicates activity, ^{d}

indicates deletion, and ^{r} indicates resistance to feedback inhibition.

7. The segment of Claim 1, wherein said pathway is a biosynthetic pathway, further comprising a *lac* operon in which the wild-type promoter sequence has been replaced with a different promoter sequence, and from which the *lac* repressor has been removed by site-specific deletion.

8. A recombinant DNA segment, comprising a methanol oxidase encoding gene and a dihydroxyacetone synthase encoding gene, both of which genes have been obtained from a methylotropic organism, and a promoter sequence directing the transcription of said genes, wherein said promoter sequence is different from the promoter sequence for said genes in said organism.

9. The DNA segment of Claim 8, further comprising a terminator site and a ribosome binding site for said genes.

10. The segment of any of Claims 1-7, wherein said segment further comprises the segment of Claim 8 or 9.

11. The segment of any of Claims 1-10 contained in a recombinant DNA vector.

12. A method for assembling a first unregulated biogenetic cassette containing DNA coding for a biosynthetic or metabolic pathway having at least two steps that is ordinarily regulated by both transcriptional regulation and feedback inhibition, wherein said cassette includes at least one gene for each step of said pathway, comprising the steps of:

obtaining a modified gene for each enzyme in said pathway that is ordinarily subject to feedback inhibition, wherein the modification in each of said modified genes destroys said inhibition;

removing the DNA sequence or sequences responsible for negative transcriptional regulation of genes for said pathway to obtain a deattenuated or derepressed operon;

assembling all of the genes coding for at least one of the enzymes for each step in said pathway, including said modified genes, into a single DNA segment;

providing a first restriction site on the 5' side of said genes and a second restriction site on the 3' side of said genes wherein said first and second restriction sites do not occur elsewhere in said segment; and

cloning said DNA segment.

13. The method of Claim 12, further comprising:

removing a promoter sequence and inserting a different promoter sequence for said genes into said segment.

14. The method of Claim 12 or 13, wherein the product of said pathway is an amino acid.

15. The method of Claim 14, wherein said amino acid is L-isoleucine.

16. The method of Claim 14, wherein said amino acid is L-valine.

17. The method of Claim 15, wherein said removing and assembling steps comprise:

(a) removing the attenuator sequence from an *ilv*GMEDA operon to form a deattenuated operon;

(b) replacing the wild-type *ilv*A with L-isoleucine-resistant *ilv*A^r; and

(c) assembling together in a single polymeric DNA molecule the deattenuated operon of step (a) containing the *ilv*A^r of step (b) and *ilv*YC to form a biogenetic cassette coding for the biosynthesis of L-isoleucine.

18. A method for creating a biogenetic cassette for methanol utilization, comprising the steps of:

obtaining methanol oxidase and dihydroxyacetone synthase encoding genes from a methylotropic organism;

removing the wild-type promoter sequence from said genes;

creating a DNA segment containing said genes and a different promoter for said genes; and

inserting said segment into a recombinant DNA transfer vector.

19. The method of Claim 18, further comprising the step of including ribosome binding and terminator sites in said segment.

20. The method of any of Claims 12-17, further comprising the step of combining said first biogenetic cassette with the biogenetic cassette of Claim 18 or 19.

21. The method of any of Claims 12-20, further comprising the step of expressing said genes in a microorganism.

International Application No: PCT/ US86 / 02440

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 26, line 16 of the description [5]

A. IDENTIFICATION OF DEPOSIT [6]

Further deposits are identified on an additional sheet [X] [6a] pABC1

Name of depositary institution [6]

American Type Culture Collection

Address of depositary institution (including postal code and country) [6]

12301 Parklawn Drive
Rockville, MD 20852 USA

Date of deposit [6]	Accession Number [6]
October 7, 1986	67,230

B. ADDITIONAL INDICATIONS [7] (leave blank if not applicable). This information is continued on a separate attached sheet [ ]

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [8] (If the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS [9] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [9] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. [ ] This sheet was received with the international application when filed (to be checked by the receiving Office)

_____
(Authorized Officer)

[ ] The date of receipt (from the applicant) by the International Bureau [10]

was

_____
(Authorized Officer)

Form PCT/RO/134 (January 1981)

(January 1985)

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 26, line 21 of the description [1]

A. IDENTIFICATION OF DEPOSIT [1]

Further deposits are identified on an additional sheet [X] [1] pABC55

Name of depositary institution [4]

American Type Culture Collection

Address of depositary institution (including postal code and country) [4]

12301 Parklawn Drive
Rockville, MD 20852 USA

Date of deposit [5]	Accession Number [6]
October 7, 1986	67,228

B. ADDITIONAL INDICATIONS [7] (leave blank if not applicable). This information is continued on a separate attached sheet [ ]

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [8] (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS [9] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [9] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. [ ] This sheet was received with the international application when filed (to be checked by the receiving Office)

_____
(Authorized Officer)

[ ] The date of receipt (from the applicant) by the International Bureau [10]

was

_____
(Authorized Officer)

(January 1985)

Form PCT/RO/134 (January 1981)

International Application No: PCT/ US86 / 02440

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 27, line 12 of the description [5]

A. IDENTIFICATION OF DEPOSIT [6]

Further deposits are identified on an additional sheet [7] ☐ CU18

Name of depositary institution [8]

American Type Culture Collection

Address of depositary institution (including postal code and country) [9]

12301 Parklawn Drive
Rockville, MD 20852 USA

Date of deposit [10]	Accession Number [11]
October 7, 1986	53,549

B. ADDITIONAL INDICATIONS [12] (leave blank if not applicable). This information is continued on a separate attached sheet ☐

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [13] (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS [14] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [15] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☐ This sheet was received with the international application when filed (to be checked by the receiving Office)

_____
(Authorized Officer)

☐ The date of receipt (from the applicant) by the International Bureau [16]

was

_____
(Authorized Officer)

(January 1985)

Form PCT/RO/134 (January 1981)

International Application No: PCT/US86/02440

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 30, line 19 of the description [5]

A. IDENTIFICATION OF DEPOSIT [6]

Further deposits are identified on an additional sheet [X] [8] pCH_4

Name of depositary institution [6]

American Type Culture Collection

Address of depositary institution (including postal code and country) [6]

12301 Parklawn Drive
Rockville, MD 20852 USA

Date of deposit [6]	Accession Number [6]
October 7, 1986	67,229

B. ADDITIONAL INDICATIONS [7] (leave blank if not applicable). This information is continued on a separate attached sheet [ ]

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [9] (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS [9] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [9] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. [ ] This sheet was received with the international application when filed (to be checked by the receiving Office)

(Authorized Officer)

[ ] The date of receipt (from the applicant) by the International Bureau [10]

was

(Authorized Officer)

(January 1985)

Form PCT/RO/134 (January 1981)

AMENDED CLAIMS

[received by the International Bureau on 21 April 1987 (21.04.87); original claims 1 and 12 amended; other claims unchanged (3 pages)]

1. A recombinant DNA segment containing genes coding for at least one enzyme for each step in the biosynthetic or metabolic pathway for a desired biological product, wherein said pathway has at least two steps, and wherein *in vivo* transcription of at least one of said genes is ordinarily subject to regulation and wherein at least one of said genes codes for an enzyme that is ordinarily subject to feedback inhibition, wherein all transcriptional regulatory sequences have been inactivated or removed from said segment, and wherein each gene in said segment that would otherwise code for a feedback-inhibited enzyme has been modified so that feedback inhibition has been removed, so that upon transcription of said segment and translation of the resulting RNA, said biosynthetic or metabolic pathway is unregulated, said segment having a first restriction site and a second restriction site with said genes located between said first and second restriction sites, wherein said first and second restriction sites do not occur elsewhere in said segment.

2. The segment of Claim 1, coding for the biosynthetic pathway for an amino acid.

3. The segment of Claim 2, wherein said amino acid is L-isoleucine.

4. The segment of Claim 2, wherein said amino acid is L-valine.

5. The segment of any of Claims 1-4, wherein said genes are in natural chromosomal order.

6. The segment of Claim 3, wherein said segment comprises *ilv*$p^{d}a^{d}G^{+}MEDA^{r}YC$, wherein p indicates a promoter sequence, e indicates a leader sequence, a indicates an attenuator sequence, $^{+}$ indicates activity, d

indicates deletion, and ^{r} indicates resistance to feedback inhibition.

7. The segment of Claim 1, wherein said pathway is a biosynthetic pathway, further comprising a *lac* operon in which the wild-type promoter sequence has been replaced with a different promoter sequence, and from which the *lac* repressor has been removed by site-specific deletion.

8. A recombinant DNA segment, comprising a methanol oxidase encoding gene and a dihydroxyacetone synthase encoding gene, both of which genes have been obtained from a methylotropic organism, and a promoter sequence directing the transcription of said genes, wherein said promoter sequence is different from the promoter sequence for said genes in said organism.

9. The DNA segment of Claim 8, further comprising a terminator site and a ribosome binding site for said genes.

10. The segment of any of Claims 1-7, wherein said segment further comprises the segment of Claim 8 or 9.

11. The segment of any of Claims 1-10 contained in a recombinant DNA vector.

12. A method for assembling a first unregulated biogenetic cassette containing DNA coding for a biosynthetic or metabolic pathway having at least two steps that is ordinarily regulated by both transcriptional regulation and feedback inhibition, wherein said cassette includes at least one gene for each step of said pathway, comprising the steps of:

obtaining a modified gene for each enzyme in said pathway that is ordinarily subject to feedback inhibition, wherein the modification in each of said modified genes destroys said inhibition;

removing the DNA sequence or sequences responsible for negative transcriptional regulation of genes for said pathway to obtain a deattenuated or derepressed operon;

assembling all of the genes coding for at least one of the enzymes for each step in said pathway, including said modified genes, into a single DNA segment;

providing a first restriction site on the 5' side of said genes and a second restriction site on the 3' side of said genes wherein said first and second restriction sites do not occur elsewhere in said segment; and

cloning said DNA segment.

13. The method of Claim 12, further comprising:

removing a promoter sequence and inserting a different promoter sequence for said genes into said segment.

14. The method of Claim 12 or 13, wherein the product of said pathway is an amino acid.

15. The method of Claim 14, wherein said amino acid is L-isoleucine.

16. The method of Claim 14, wherein said amino acid is L-valine.

17. The method of Claim 15, wherein said removing and assembling steps comprise:

(a) removing the attenuator sequence from an *ilv*GMEDA operon to form a deattenuated operon;

(b) replacing the wild-type *ilv*A with L-isoleucine-resistant *ilv*A^r; and

(c) assembling together in a single polymeric DNA molecule the deattenuated operon of step (a) containing the *ilv*A^r of step (b) and *ilv*YC to form a biogenetic cassette coding for the biosynthesis of L-isoleucine.

18. A method for creating a biogenetic cassette for methanol utilization, comprising the steps of:

obtaining methanol oxidase and dihydroxyacetone synthase encoding genes from a methylotropic organism;

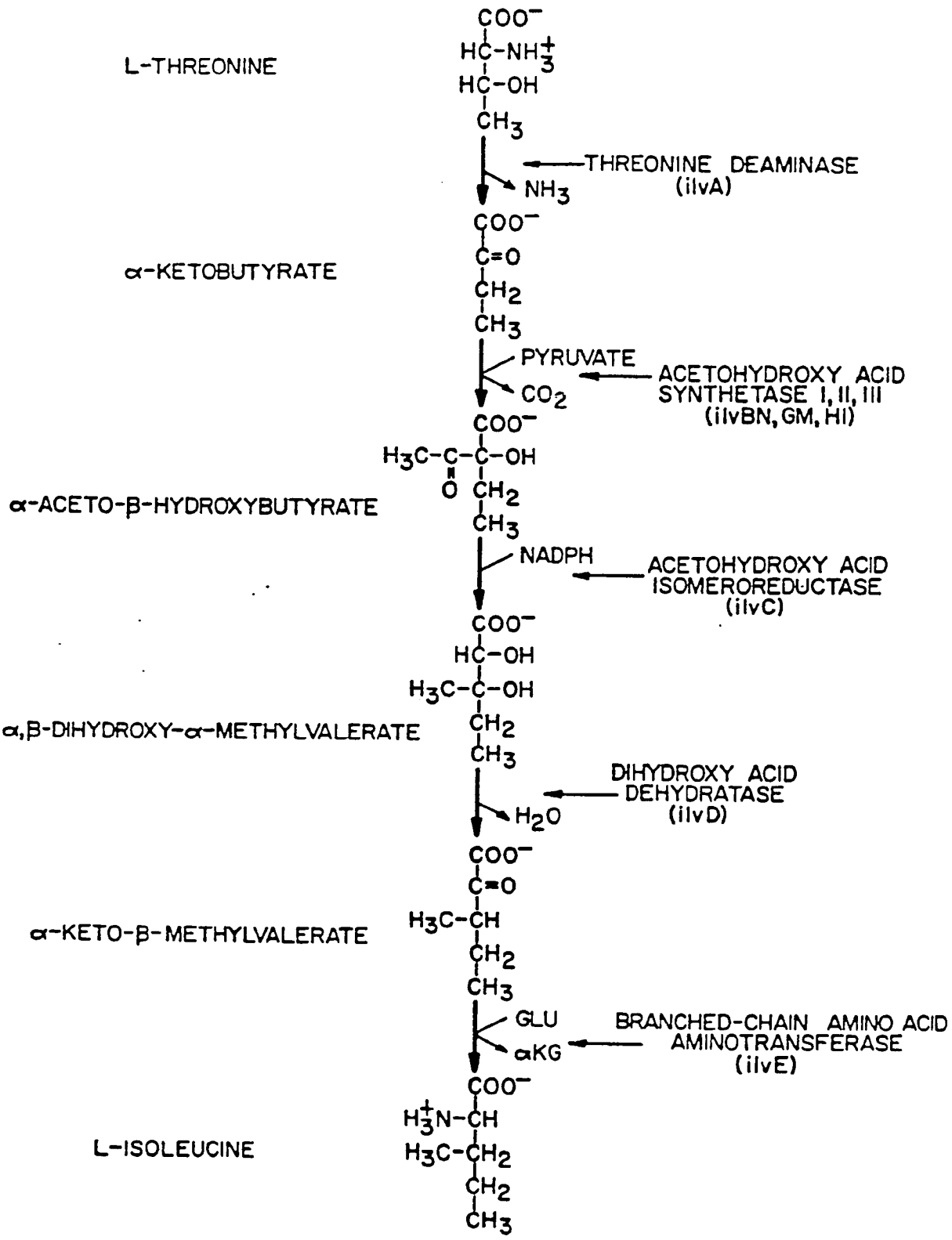
L-THREONINE
THREONINE DEAMINASE
(ilvA)
NH3
α-KETOBUTYRATE
PYRUVATE
CO2
ACETOHYDROXY ACID
SYNTHETASE I, II, III
(ilvBN, GM, HI)
α-ACETO-β-HYDROXYBUTYRATE
NADPH
ACETOHYDROXY ACID
ISOMEROREDUCTASE
(ilvC)
α,β-DIHYDROXY-α-METHYLVALERATE
DIHYDROXY ACID
DEHYDRATASE
(ilvD)
H2O
α-KETO-β-METHYLVALERATE
GLU
αKG
BRANCHED-CHAIN AMINO ACID
AMINOTRANSFERASE
(ilvE)
L-ISOLEUCINE

FIG. 1

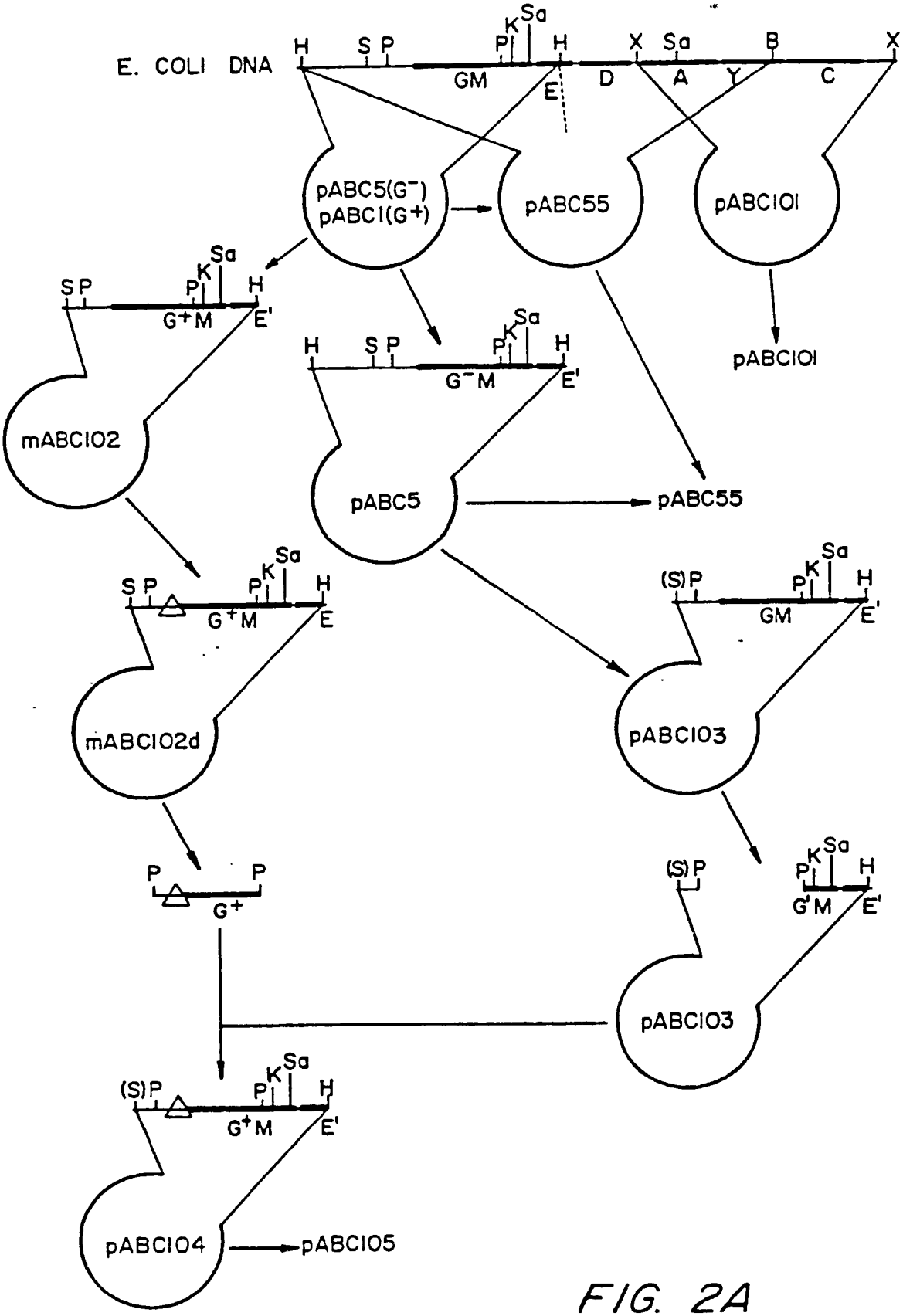
E. COLI DNA
H S P P K Sa H X Sa B X
GM E' D A Y C
pABC5(G⁻)
pABC1(G+)
pABC55
pABC101
pABC101
mABC102
pABC5
pABC55
mABC102d
pABC103
pABC103
pABC104
pABC105

FIG. 2A

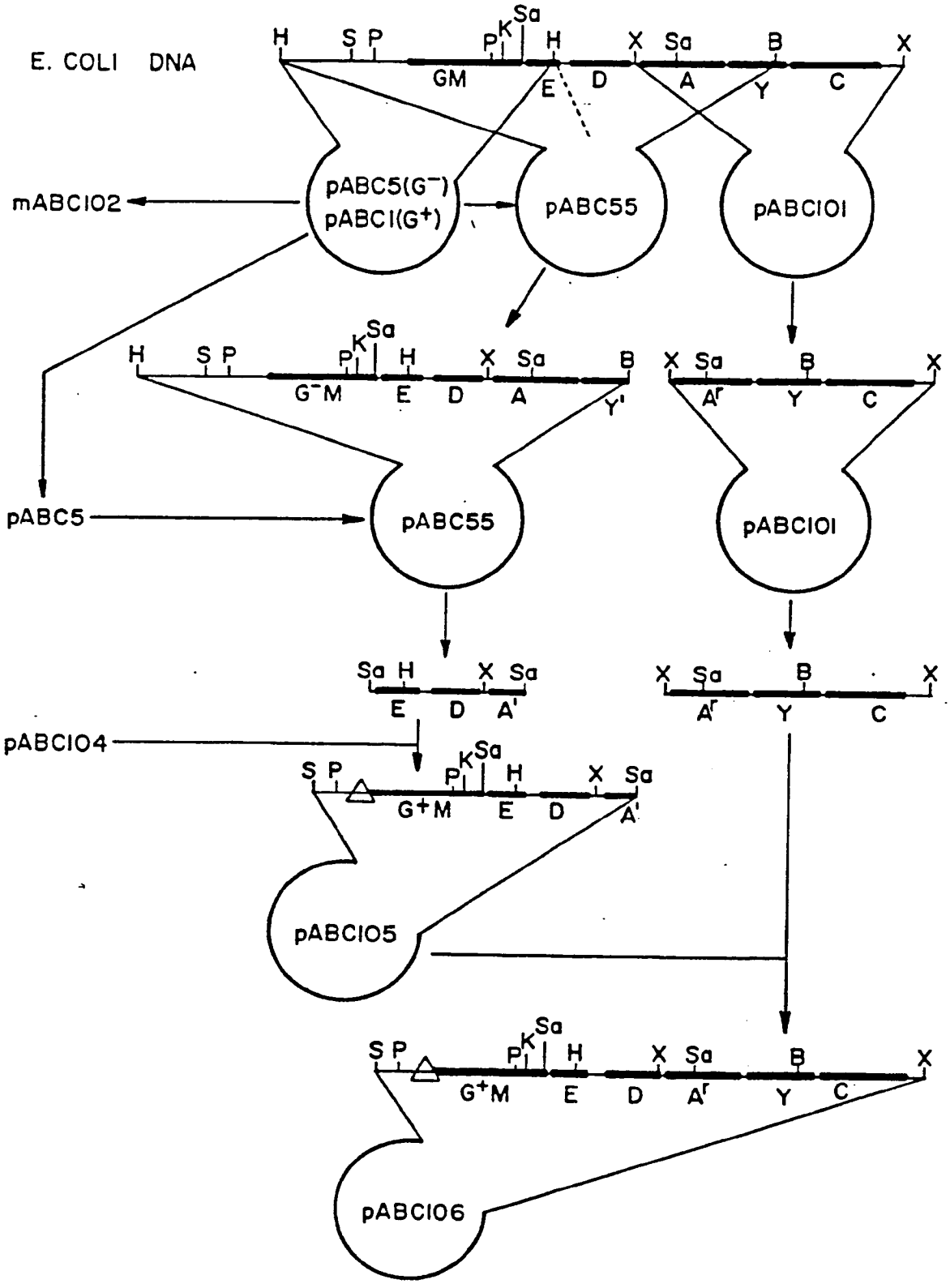
E. COLI DNA
H S P P K Sa H X Sa B X
GM E D A Y C
mABC102
pABC5(G⁻)
pABC1(G⁺)
pABC55
pABC101
H S P P K Sa H X Sa B
G⁻M E D A Y'
X Sa B X
Aʳ Y C
pABC5
pABC55
pABC101
Sa H X Sa
E D A'
X Sa B X
Aʳ Y C
pABC104
S P P K Sa H X Sa
G⁺M E D A'
pABC105
S P P K Sa H X Sa B X
G⁺M E D Aʳ Y C
pABC106

FIG. 2B

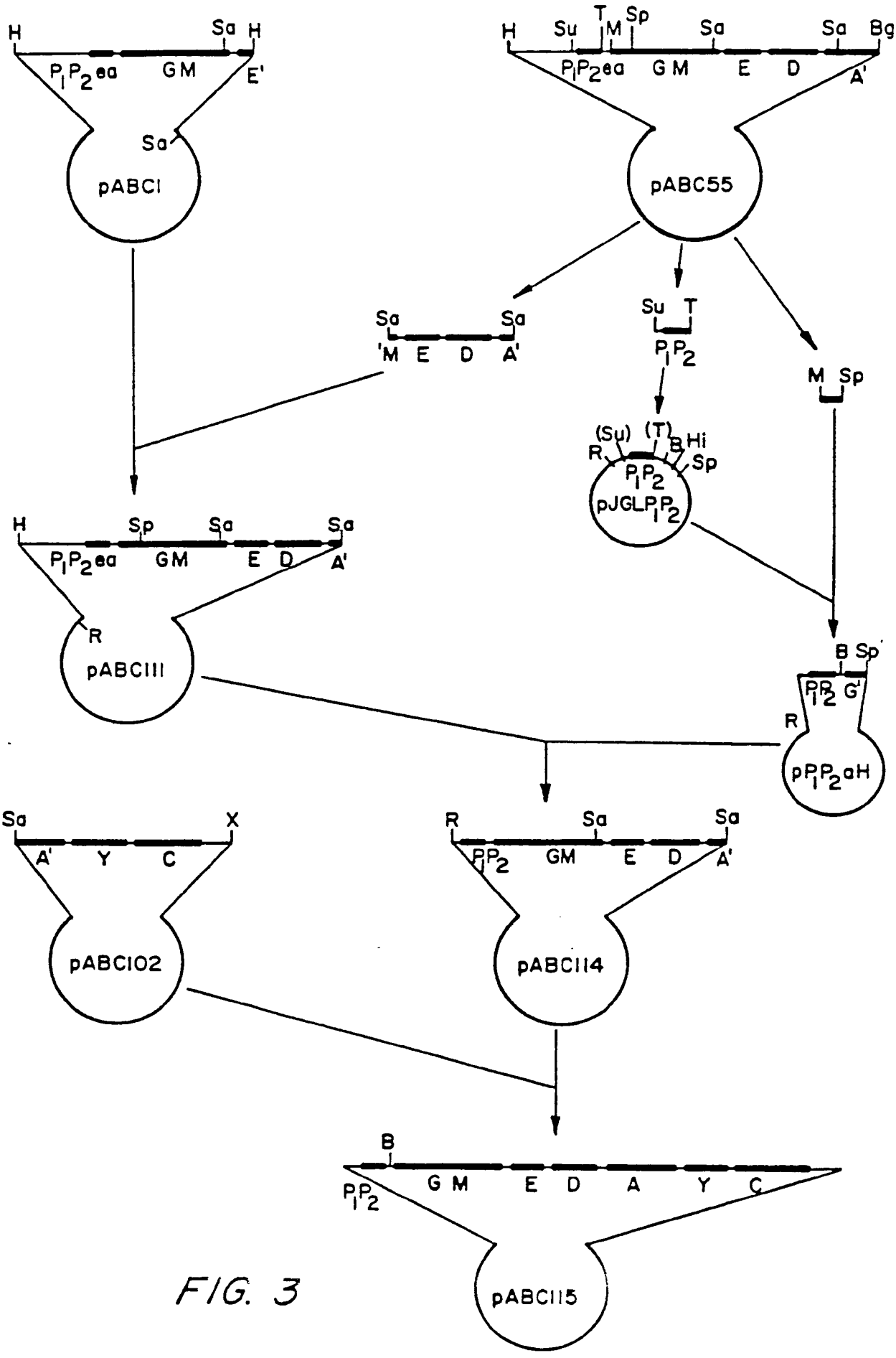
H
Sa H
P1P2 ea
GM
E'
Sa
pABC1
H Su T M Sp Sa Sa Bg
P1P2 ea GM E D A'
pABC55
Sa Sa
'M E D A'
Su T
P1P2
M Sp
(Su) (T) B Hi
R
Sp
P1P2
pJGLP1P2
H Sp Sa Sa
P1P2 ea GM E D A'
R
pABC111
B Sp
P1P2 G'
R
pP1P2aH
Sa X
A' Y C
pABC102
R Sa Sa
P1P2 GM E D A'
pABC114
B
P1P2
GM E D A Y C
pABC115

FIG. 3

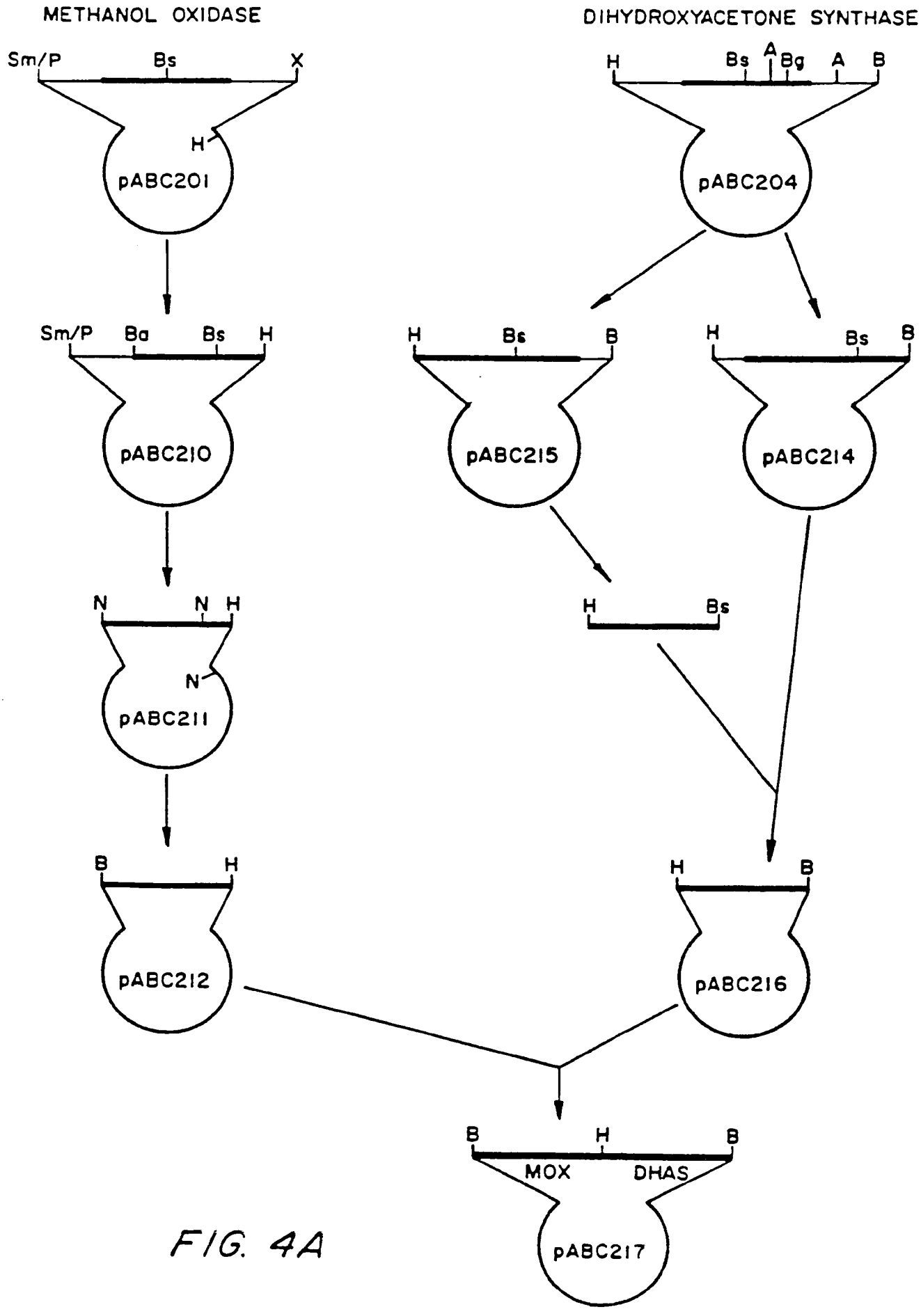
METHANOL OXIDASE
DIHYDROXYACETONE SYNTHASE
Sm/P
Bs
X
H
pABC201
H
Bs
A
Bg
A
B
pABC204
Sm/P
Ba
Bs
H
pABC210
H
Bs
B
pABC215
H
Bs
B
pABC214
N
N
H
N
pABC211
H
Bs
B
H
pABC212
H
B
pABC216
B
H
B
MOX
DHAS
pABC217

FIG. 4A

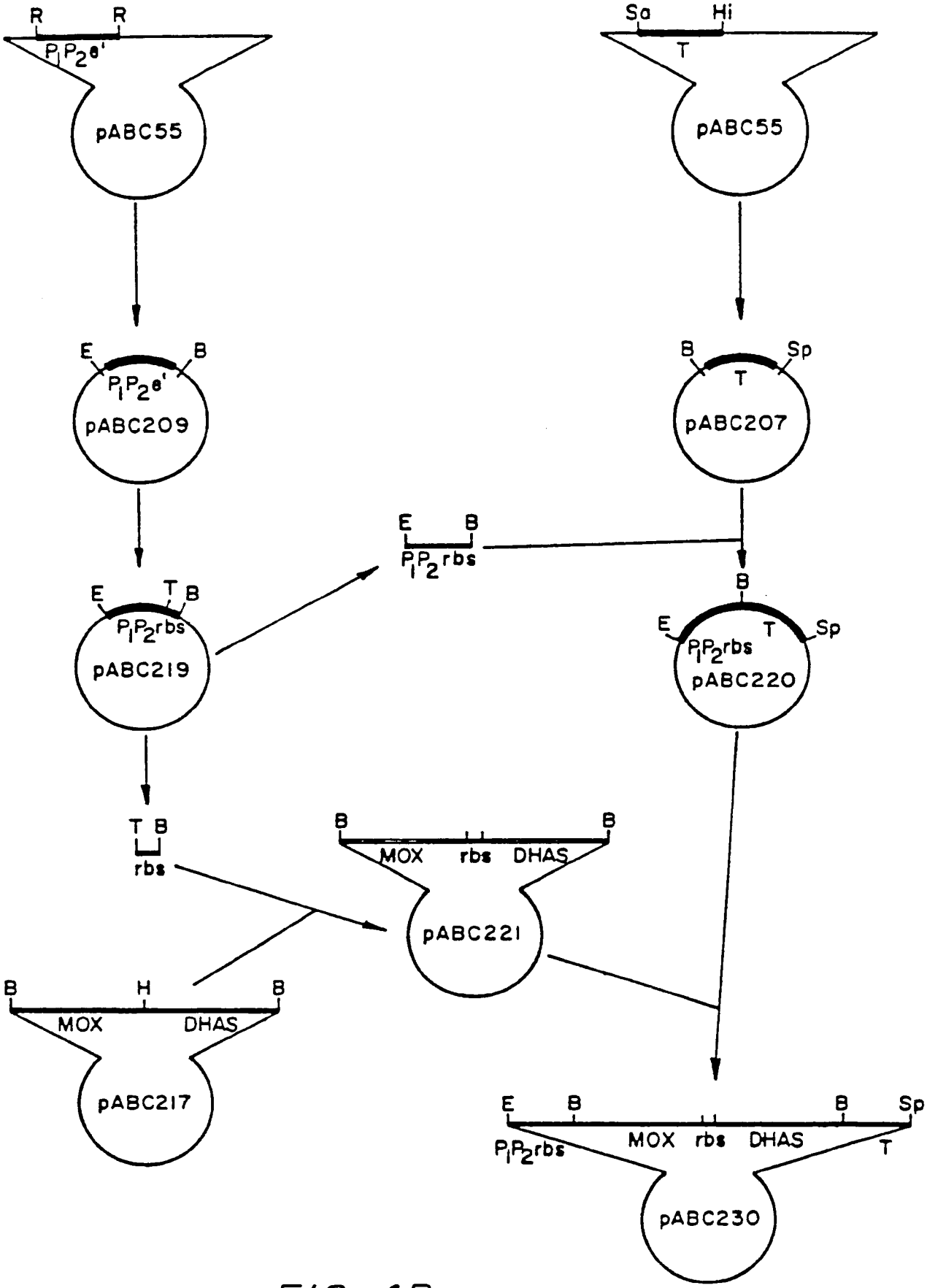
R
R
P1P2e'
pABC55
Sa
Hi
T
pABC55
E
B
P1P2e'
pABC209
B
Sp
T
pABC207
E
B
P1P2rbs
E
T
B
P1P2rbs
pABC219
B
E
T
Sp
P1P2rbs
pABC220
T
B
rbs
B
MOX
rbs
DHAS
B
pABC221
B
H
B
MOX
DHAS
pABC217
E
B
B
Sp
P1P2rbs
MOX
rbs
DHAS
T
pABC230

FIG. 4B

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 86/02440

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (4): C07H 21/04; C12N 15/00

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System	Classification Symbols
U.S.	435/68,70,71,91,106,107,108,109,110,113,114,115, 116,172.1,172.3,253,255,256,317;536/27,935/6,10, 14,16,23,29,39,40,60,61,72,73

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [5]

CA SEARCH DATA BASE: 1967-1987

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
Y	US, A, 4,278,765 (DEBABOV ET AL) 14 July 1981. See the entire document.	1-9 & 12-19
X/Y, P	US, A, 4,371,614 (ANDERSON ET AL) 1 February 1986. See the entire document.	1,2,5,12, 13 & 14 / 3,4,6-9 & 15-19
Y	Chemical Abstracts, Volume 102, No. 21, issued 27 May 1985 (Columbus, Ohio, USA), CHI ET AL, "In vivo cloning of E. coli K-12 trpL (Δatt) $trpE^{FBR}$ gene using the plasmid RP_4:: Mucts 61", see page 154, the abstract bridging columns 1-2, the Abstract No, 180175s, Han'guk Saenghwa Hakhoechi 1984, 17(3) 266-74 (Korean).	1-9 & 12-19

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
04 February 1987	18 FEB 1987
International Searching Authority [1]	Signature of Authorized Officer [20]
ISA/US	James Martinell

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No [18]
Y	A.M. LEDBOER ET AL, "Molecular cloning and characterization of a gene coding for methanol oxidase in Hansenula polymorpha", Nucleic Acids Research, Volume 13, Number 9, pages 3063-3082, published 10 May 1985 by IRL Press Limited (Oxford, UK), see especially abstract and pages 3072, 3073, 3075, 3078, 3079, and 3080.	1-9 & 12-19
Y	Z.A. JANOWICZ ET AL, "Cloning and characterization of the DAS gene encoding the major methanol assimilatory enzyme from the methylotrophic yeast Hansenula polymorpha", Nucleic Acids Research, Volume 13, Number 9, pages 3043-3062, published 10 May 1985 by IRL Press Limited (Oxford, UK), see especially abstract and pages 3055-3061.	1-9 & 12-19
Y	D.H. CALHOUN ET AL, "Cloning of the ilvA538 gene coding for the feedback-hypersensitive threonine deaminase from Escherichia coli K-12", Journal of Bacteriology, Volume 151, Number 1, pages 274-280, published July 1982 by the American Society for Microbiology (Washington, D.C., USA), see abstract and page 279.	1-9 & 12-19
Y	R.P. LAWTHER ET AL, "Molecular basis for valine resistance in Escherichia coli K-12", Proceedings of the National Academy of Science, Volume 78, Number 2, pages 922-925, published February 1981 by the National Academy of Science (Washington, D.C., USA), see entire document.	1-9 & 12-19
Y	R.P. LAWTHER ET AL, "Multivalent translational control of transcription termination at attenuator of ilvGEDA operon of Escherichia coli K-12", Proceedings of the National Academy of Science, Volume 77, Number 4, pages 1862-1866, published April 1980 by the National Academy of Science (Washington, D.C., USA), see entire document.	1-9 & 12-19

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No [18]
Y	R.P. LAWTHER ET AL, "DNA sequence fine-structure analysis of ilvG(IlvG+) mutations of Escherichia coli K-12", Journal of Bacteriology, Volume 149, Number 1, pages 294-298, published January 1982 by The American Society for Microbiology (Washington, D.C., USA), see entire document.	1-9 & 12-19
Y	C.A. HAUSER ET AL, "Nucleotide sequence of the ilvB multivalent attenuator region of Escherichia coli K-12", Nucleic Acids Research, Volume 11, Number 1, pages 127-139, published 11 January 1983 by IRL Press Limited (Oxford, UK), see entire document.	1-9 & 12-19
Y	J.M. SMITH ET AL, "Polarity and the regulation of the ilv gene cluster in Escherichia coli strain K-12", Molecular and General Genetics, Volume 148, pages 111-124, published 1976 by Springer-Verlag (Berling, FRG), see especially abstract, Figure 1 and its legend, and pages 119-123.	1-9 & 12-19
X,P Y,P	EP, A, 0,190,921 (ENGENICS, INC.), 13 August 1986, see entire document.	1,2,4, 5,12,13, 14, & 16 3,6-9, 15, & 16-19
X Y	EP, A, 0,145,156 (BIOTECHNICA INTERNATIONAL, INC.), 19 June 1985, see entire document.	1,2,5,12, 13, & 14 3,4,6-9 & 15-19

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

Y	Chemical Abstracts, Volume 99, No. 25, issued 19 December 1983, (Columbus, Ohio, USA), CHI ET AL, "Construction of Escherichia coli K-12 trpL (Δatt) $trpE^{FBR}$ mutant and its characterization", see page 176, column 1, the Abstract No. 207208p, Han'guk Saenghwa Hakhoechi 1983, 16(3), 249-54 (Korean).	1-9 & 12-19

V. [X] OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [10]

This International search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. [ ] Claim numbers ....., because they relate to subject matter [12] not required to be searched by this Authority, namely:

2. [X] Claim numbers 10, 11, 20 and 21, because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [13], specifically:

These claims are unsearchable because they are improper multiple dependent claims because they do not refer to claims from which they depend in the alternative only and because a multiple dependent claim shall not serve as the basis for any other multiple dependent claim (see PCT Rule 6.4).

VI. [ ] OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [11]

This International Searching Authority found multiple inventions in this international application as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. [ ] As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

[ ] The additional search fees were accompanied by applicant's protest.

[ ] No protest accompanied the payment of additional search fees.